CODE APERTURE IMAGING FROM THE VISIBLE TO X-RAY

by

Edgar Eduardo Salazar

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Spring 2022

CODE APERTURE IMAGING FROM THE VISIBLE TO X-RAY

by

Edgar Eduardo Salazar

Approved: ____________________
Jamie Phillips, Ph.D.
Chair of Electrical and Computer Engineering

Approved: ____________________
Levi T. Thompson, Ph.D.
Dean of the College of Engineering

Approved: ____________________
Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____
Gonzalo Arce, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____
Michael Piovoso, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____
Javier Garcia-Frias, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____
Kenneth Barner, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Alejandro Parada-Mayorga, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

I would like to thank Dr. Gonzalo Arce for giving me the opportunity to work with him on my PhD. His guidance has always been appreciated. I also want to recognize the fundamental role of Dr. Alejandro Parada in my PhD; his advise helped me succeed during my studies. My sincere gratitude to Dr. Michael Piovoso, whose passion for teaching inspired me to follow a path in academia. I also want to express my gratitude to Dr. Javier Garcia-Frias and Dr. Kenneth Barner, for serving on my PhD committee; their feedback and ideas have been invaluable.

This journey would have not been possible without the help and support of my ECE colleagues and all the guys at the computational imaging and data science group; our formal and informal discussions were of great help in shaping my ideas.

To my mother and my brother in Colombia, I want to say thanks; they were always there no matter the circumstances. To my extended family in Bolivia, many thanks; your constant support during this time was comforting.

To my beloved wife, Yara, the love of my life, whose courage inspires me every day, and to my beautiful daughter, Noita, the owner of my heart, whose smile makes me feel like the luckiest man on earth. I just want to say that everything I am and will be, is because and thanks to you.

Finally, to my Father, Alfonso: I remember you every day.

TABLE OF CONTENTS

Appendix

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Inverse problems have become a topic of broad interest in science and engineering, given that only incomplete data might be available as a consequence of hardware limitations that can be intentionally or unintentionally created. Spectral image acquisition has been positively favored from inverse problems, such that conventional sensing methods present limitations linked to the time required to fully acquire a hyperspectral scene. By implementing an optical device that accounts for pseudo-random compressed sampling, also known as coded aperture, inverse problems algorithms may be run and a fully hyperspectral scene is reconstructed from few 2-dimensional projections.

The *Coded Aperture Snapshot Spectral Imager* (CASSI), represents the first attempt to experimentally validate the recovery of spectral information from incomplete data. Although results in terms of spatial and spectral quality are remarkable, CASSI possesses limitations which inherently affect the reconstructed datacubes. One of them, is the fact that a spectrally dispersed image must arrive at the detector in order to estimate the spectral information. *The Spatial Spectral Compressive Spectral Imager* (SSCSI) was proposed as an alternative to overcome, from the hardware point of view, the many drawbacks of CASSI. Nevertheless, when first proposed, the resolution limits of the SSCSI were unknown, nor was a discrete sensing model well understood. This dissertation deeply analyses the physics and mathematics of the coding and information capturing process in SSCSI. A strong theoretical analysis of the continuous model is developed, and the proposed hypothesis is validated through a test-bed implementation. As it was done for CASSI and other compressive spectral imagers (CSI), the optimization of the implemented coding patterns was further explored, based on the developed sensing model. The *Restricted Isometry Property*, a sensing matrix metric widely utilized in Compressive Sampling, was adapted to the SSCSI framework. Results show that the optimal codes must exhibit certain structure that can be explained using the blue-green noise patterns.

This work then further explores coded aperture imagers in the X-rays regime, where we focus on Compton backscatter X-rays sensing, a commonly used technique to identify low atomic number and organic materials. Unlike traditional X-ray backscatter imaging devices, the proposed work avoids pencil-beam X-rays illumination. Instead, we adopt cone-beam coded illumination to capture the information parallel-wise. The proposed architecture, coined *Compressive X-rays Compton Backscattering Imager*, is analyzed in detail, and a discrete measurements model is proposed and tested in simulations using *Geant4 Application for Tomographic Emission* (GATE), under realistic conditions for several types of targets.

Chapter 1

INTRODUCTION

Coded aperture imaging is not a new technique. It was first proposed by Dicke in [1], in the context of γ-rays, as an alternative to increase the Signal-to-Noise Ratio (SNR) of the captured images while preserving the angular resolution of a pinhole camera. In the context of imaging inverse problems, coded apertures are used to control the modulation of the information prior to the sensing stage, such that the vectorized scene $\mathbf{x} \in \mathbb{R}^n$ can be recovered from the vector of measurements $\mathbf{g} \in \mathbb{R}^m$, defined as follows:

$$\mathbf{g} = \mathbf{H}\mathbf{x} + \epsilon, \tag{1.1}$$

where ϵ is the intrinsic signal independent noise of the capturing process, and $\mathbf{H} \in \mathbb{R}^{m \times n}$ is the sensing matrix that models the physics of the sensing and coding process in the imager. Given that $m < n$, this is considered and ill-conditioned problem. Different approaches have been explored to solve Eq. (1.1). One of the first proposed was through the *Modified Uniformly Redundant Arrays* patterns (MURA) [2], which are specifically designed such that $\hat{\mathbf{H}}_{MURA} \times \mathbf{H}_{MURA} \approx \mathbf{I}$, where $\mathbf{H}_{MURA}$ and $\hat{\mathbf{H}}_{MURA}$ are the MURA and complementary MURA sensing matrices, and $\mathbf{I} \in \mathbb{R} \in n \times n$ is the identity matrix. Other approaches find $\mathbf{x}$ as the solution of an optimization problem as follows:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} ||\mathbf{g} - \mathbf{H}\mathbf{x}||_2^2 + \tau R(\mathbf{x}), \tag{1.2}$$

where the first element of the equation is known as the conformity term, which restricts the solution to be in accordance with the measured data, while the second element represents prior knowledge of the scene to be recovered; the variable τ controls the influence of such prior knowledge in the final solution of the optimization problem. A commonly used assumption is data sparsity on a given domain, such that Eq. (1.2) can be rewritten as shown below:

$$\hat{\boldsymbol{\pi}} = \min_{\boldsymbol{\pi}} ||\mathbf{g} - \mathbf{H}\boldsymbol{\Psi}\boldsymbol{\pi}||_2^2 + \tau||\boldsymbol{\pi}||_1, \tag{1.3}$$

where $\mathbf{x} = \boldsymbol{\Psi}\boldsymbol{\pi}$, and $\boldsymbol{\Psi}$ is a transformation basis that allows the scene to be represented using a sparse vector $\boldsymbol{\pi}$ with few non-zero coefficients This concept represents the core of Compressive Sensing [3].

1.1 Compressive spectral imaging

Hyperspectral Imaging is a technique where spatial and spectral information from a scene are simultaneously registered. It has applications in several areas, such as medical imaging [4], remote sensing [5], artwork conservation [6], and food quality [7], among others. Conventional acquisition methods, such as push-broom or whisk-broom, although well-defined techniques, have several disadvantages, such as the needed time to run the scanning process, or the massive amount of data to be stored[8]. The *Coded Aperture Snapshot Spectral Imager* (CASSI), proposed by Wagadarikar et. al. [9], represents the first real attempt to couple inverse problems and compressive sensing to spectral imaging. Unlike the aforementioned methods, CASSI allows the acquisition of a full spectral scene through few coded 2-dimensional measurements together with the solution of Eq. (1.2). Figure 1.1 [10] shows the main components of CASSI. From its advent, the performance of CASSI has been optimized through several hardware modifications. One of the first ones was proposed by Wu. et. al. [11], where the basic lithographic mask was replaced by a programmable Digital Micromirror Device (DMD) in order to easily implement and change the patterns. Arguello and Arce [12], replaced the blocking-unblocking coded aperture masks with pixelated optical filters such that the coding process was enriched; this architecture, known as Colored-CASSI, represents

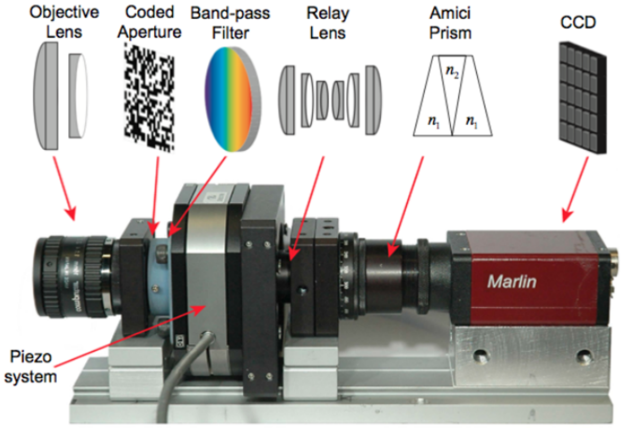

Figure 1.1: Coded Aperture Snapshot Spectral Imaging hardware components. Three main components can be mentioned: the coded aperture, the spectral disperser (prism), and the gray-scale sensor.

a major improvement in the final quality of the reconstruction results. Rueda et. al. introduced the concept of *Higher Order Discretization Model*, where the spectral dispersion of the scene was seen as singly-independent shearing for each voxel component [13], such that the sensing model resembles the actual physics and light propagation. Other improvements on CASSI and CSI include the use of different sources of side information in the sensing stage [14, 15] and the location of a colored filter mask on top of the sensor [16]. Lin et. al. [17] proposed the *Spatial Spectral Compressive Spectral Imager* (SSCSI), as an alternative to the Dual-Disperser CASSI [18]. In the SSCSI, the spatial-spectral coding of the scene is enriched by tuning the coded aperture position inside the imager, while the scene arrives totally in-focus at the sensor. Improvements in terms of quality of the scene are considerable. Nevertheless, a theoretical analysis of the SSCSI in terms of spatial and spectral resolution is not available in the literature. This limits the real applications of SSCSI, such that a discrete measurements model that considers the real quantized data has not yet been proposed.

1.2 Compton backscattering imaging

Compton Backscattering imaging (CBI) is a technique widely implemented to detect low atomic and organic materials in a given body. It is implemented for under-vehicle and airport luggage inspection and screening of buildings, among other uses [19, 20, 21]. CBI is considered a single-side technique, given that source and detector are located on the same side of the setup. The state-of-the-art acquisition technology for CBI can be classified into two different approaches [22]. In the first one, the information is acquired through a pixel-by-pixel scanning, using the so-called flying spot technique, which consists of a fan beam collimator and a chopper wheel whose rotation accounts for the scanning in one axis, while the relative movement between the source and the body accounts for the scanning in the remaining axis. Figure 1.2 shows the basic functionality of this method. In the second approach, the body is fully bathed by a cone beam and the backscattered photons are focused on a 2-dimensional detector using proper optical imagery, such as the lobster eye [23], a coded aperture [24], or Fresnel plates [25]. Two main factors are crucial in CBI and any ionizing radiation-based human scanning. The first one is the radiation dose, as accumulated exposure over years might cause malignant mutations that can lead to certain types of cancer [26, 27]; the second one is the time of acquisition, given that security checkpoints always intend to detect hazardous objects in real time. Although compressive sensing and inverse problems algorithms have been used as a tool to solve similar problems in different contexts such as CSI, CBI has not yet been conceived through this framework.

1.3 Motivation

The increasing interest of coded aperture imaging in inverse problems has led to the continuous improvement of reconstruction algorithms as well as the implemented hardware. Today, CSI is a well-established research area that is rapidly advancing in terms of final image quality with less sensed data, and most of the proposed cameras in the literature depend on different types of coding techniques to randomly sample

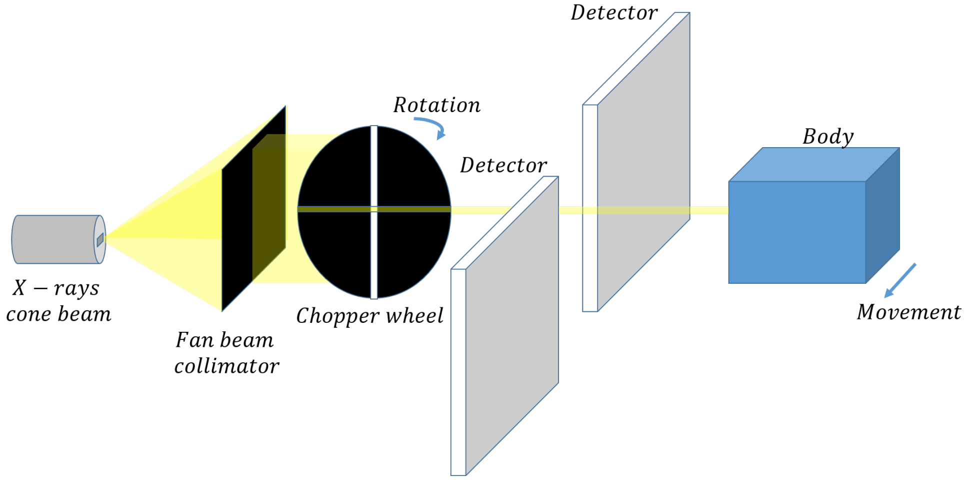

Figure 1.2: Flying spot architecture. An X-ray cone beam source passes through a fan beam collimator and then through a chopper wheel which acts as rotating pencil beam collimator; this allows the body to be vertically scanned. In order to fully scan the body, a horizontal displacement must exist.

the data. This dissertation is motivated by the need to understand the role of the coded aperture in two different scenarios; the first one is CSI, and in particular the case of SSCSI. A rigorous analysis of the spatial and spectral resolution limits is done in theory and experiments, to then formulate a closed expression of the captured discrete measurements. Subsequently, and using this model, an optimization of the SSCSI coding patterns is proposed.

The second scenario is CBI; here, the CBI sensing process is done through coded illumination, and the image is recovered through inverse problems algorithms. Given that this is the first work of its kind, a discrete measurements model is proposed and tested in simulations and experiments to evaluate its viability.

1.4 Dissertation organization

This dissertation is organized in three different parts. In the first part, a detailed analysis of the SSCSI resolution limits is done, and a discrete measurements model is proposed. In the second part, the optimization of the SSCSI coded aperture patterns is developed. Finally, the third part focuses on CBI seen through an inverse problem framework.

1.5 Research contributions

The research contributions in terms of publications are itemized below and classified in conferences and journals.

1.5.0.1 Conferences

- E. Salazar A. Parada, and G. R. Arce. *Spectral zooming in SSCSI compressive spectral imagers.* Computational Optical Sensing and Imaging. Optical Society of America, 2018.
- E. Salazar, A. Parada, and G. R. Arce. *Spatial super-resolution reconstruction via SSCSI compressive spectral imagers.* Computational Optical Sensing and Imaging. Optical Society of America, 2018.
- E. Salazar, and G. R. Arce. *Optimal coding patterns in spatial spectral compressive spectral imagers.* Propagation Through and Characterization of Atmospheric and Oceanic Phenomena. Optical Society of America, 2019.
- E. Salazar and G. R. Arce. *On the Move Compton Backscattering Scanning.* SPIE-Defense and Commercial Sensing. 2022 (Invited paper).

1.5.0.2 Journals

- E. Salazar, A. Parada, and G. R. Arce. *Spectral zooming and resolution limits of spatial spectral compressive spectral imagers.* IEEE Transactions on Computational Imaging 5.2 (2019): 165-179.
- E. Salazar, and G. R. Arce. *Coded aperture optimization in spatial spectral compressive spectral imagers.* IEEE Transactions on Computational Imaging 6 (2020): 764-777.
- E. Salazar, and G. R. Arce. *X-ray Compton Backscattering Imaging via Structured Light.* Accepted for publication in Optics Express.

Chapter 2

THE SPATIAL SPECTRAL COMPRESSIVE SPECTRAL IMAGER

2.1 Introduction

CSI allows the acquisition of 3-dimensional data (2 spatial dimensions, one spectral dimension), through 2-dimensional random projections. The CASSI relies on three basic hardware components; the first one is the coded aperture, which modulates the information; the second one is the prism, which introduces spectral dispersion to the scene; and the last one is the 2-dimensional gray-scale sensor, which captures the coded light (see Fig. 1.1). Although CASSI offers remarkable results in terms of spatial and spectral quality, it is well known that it has several drawbacks, such as the previously mentioned dispersion that impairs the final spatial quality. The *Spatial Spectral Compressive Spectral Imager*, proposed by Lin et. al. [17] arose as a feasible hardware solution and it is now included in the state-of-the-art CSI architectures [28]. Unlike CASSI, in the SSCSI, the scene arriving at the sensor is not required to be spectrally dispersed, while the coding process is controlled through positioning the coded aperture at a given distance with respect to the sensor. This can be clearly seen in Fig. 2.1, where the ray-tracing diagram of both the SSCSI and CASSI are depicted. Although widely accepted, the SSCSI does not possess any formal analysis and no technical details of the sensing process are available in the literature; this is a fundamental step for future SSCSI research and commercial purposes. That is why in this chapter, a detailed analysis of the SSCSI resolution limits is developed, as a function of the coded aperture and detector physical dimensions, the coded aperture position with respect to the sensor, and the dispersion process introduced by the diffraction grating. To that end, this chapter is organized as follows: in Section II, the SSCSI sensing phenomena is

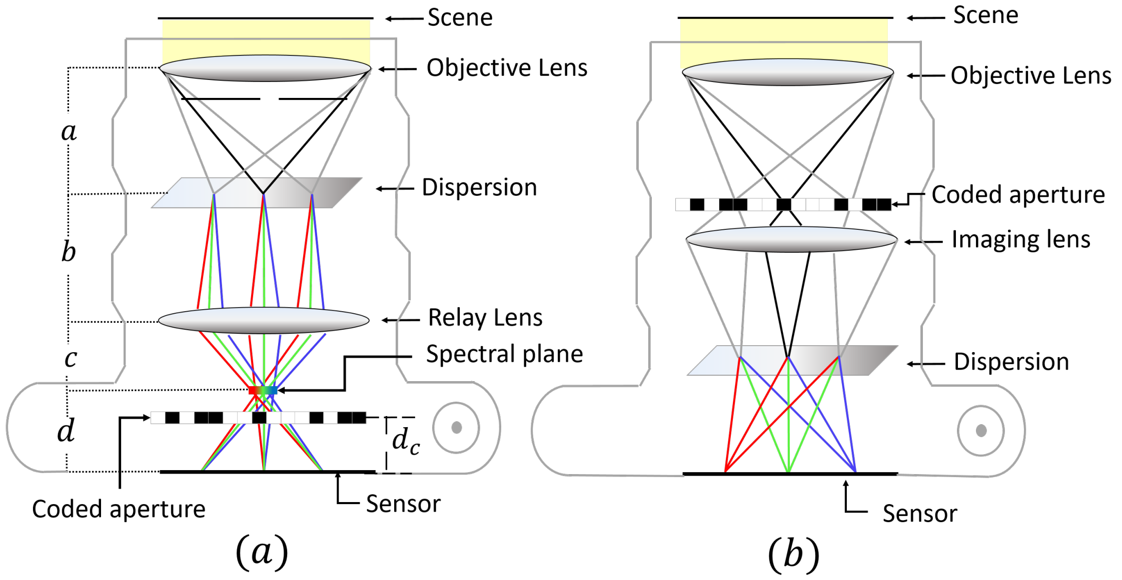

Figure 2.1: Ray tracing diagram for (a) SSCSI, where the coded aperture is placed between the spectral plane and the sensor and (b) Single Disperser CASSI, where the coded aperture is located before the spectral dispersion.

analyzed; in Section III, the spatial and spectral resolution are established, and several discrete measurements models are proposed; in Section IV, simulations are run and the proposed models are compared against conventional and Colored-CASSI; in Section V, experimental measurements are shown; finally, this chapter concludes with a section highlighting the main contributions to this research topic.

2.2 SSCSI sensing phenomena

The SSCSI optical architecture depicted in Fig. 2.1(a), makes use of an objective lens to locate an image of the scene directly on the diffraction grating, where the spectral dispersion takes place. The relay lens is then used to get an in-focus image onto the sensor. Figure 2.2 depicts the ray propagation of the scene into the sensor. Notice in Fig. 2.2 that there exists a physical position where all rays from a same wavelength converge to a single line creating a rainbow; this is known as the spectral plane, and it is extensively studied by Mohan et. al. in [29]. The width of that spectral plane, R_1, is given by the following expression

$$R_1 = \frac{bd}{b+d}\theta, \tag{2.1}$$

where b is the distance between the grating and the relay lens and d is the distance between the spectral plane and the sensor (see Fig. 2.1(a)). The variable θ is the grating dispersion angle.

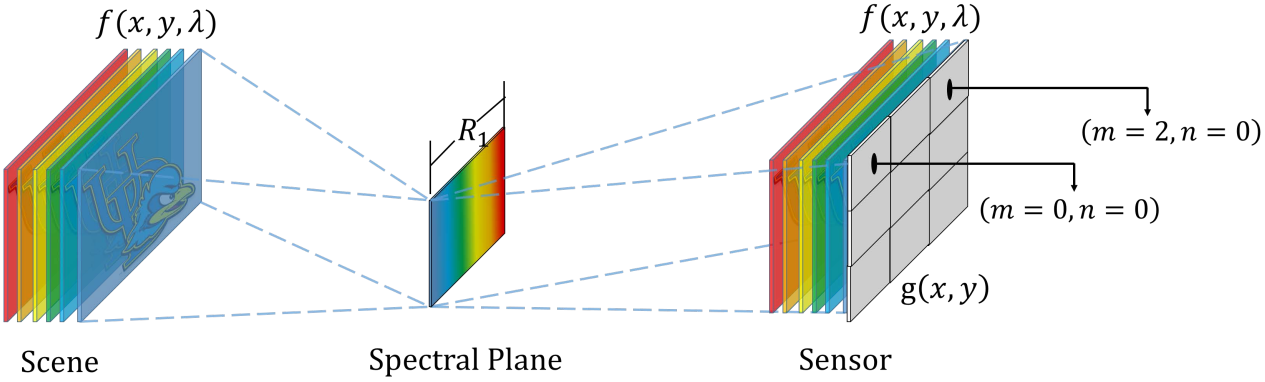

Figure 2.2: Ray propagation of the spectral scene onto the spectral plane and into the sensor. Here, $g(x, y)$ is the continuous model representation of the information captured at the sensor. The elements $(m = 0, n = 0)$ and $(m = 2, n = 0)$ of the detector array are pointed out for clarification purposes.

The coding process happening at the SSCSI is strongly influenced by the coded aperture position with respect to the sensor; this is quantified through the normalized variable $s = d_c/d$ (see Fig. 2.1), where $s = 0$ when the coded aperture is located on top of the sensor, and $s = 1$ when it is on the spectral plane. As depicted in Fig. 2.3, when $s = 1$, each resolvable spectral band is coded by a different column of the coded aperture, and a stretched version (elongated over the horizontal axis) of the pattern is seen at the sensor; as the mask moves towards the sensor, each spectral band is coded by a bigger portion of the coded aperture, and the stretching factor decreases. When $s = 0$, all spectral bands are coded by exactly the same portion of the coded aperture, and no stretching is observed at the detector. The gray-scale measurements

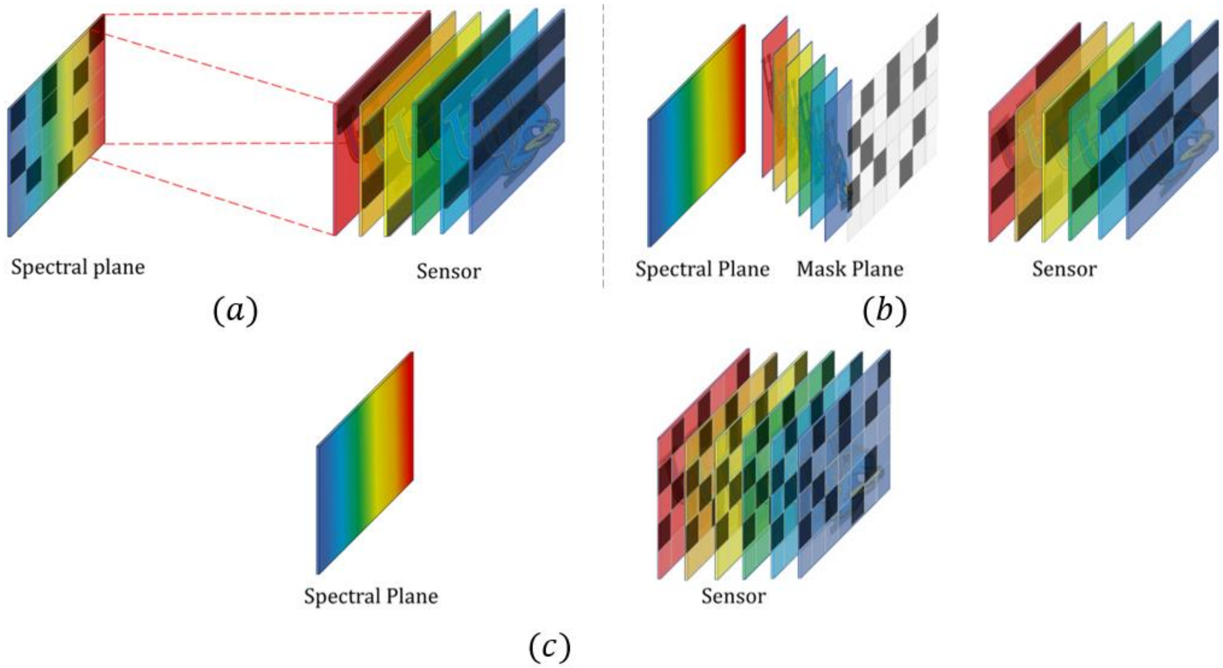

Figure 2.3: (a) Coding at the spectral plane ($s = 1$), and the coded field at the detector. (b) Coding at $0 < s < 1$ and the coded field at the detector. (c) Coding when the coded aperture is located at the detector ($s = 0$).

for different values of s can be seen in Fig. 2.4. The SSCSI sensing phenomena is described through the following expression:

$$g(x, y) = \int_{\Lambda} T(x(1 - s) + s\alpha\lambda, y) f(x, y, \lambda) d\lambda, \tag{2.2}$$

where $g(x, y)$ is the continuous model representation of the information captured at the sensor (see Fig. 2.2), $T(x, y)$ is the coded aperture function, $f(x, y, \lambda)$ is the continuous representation of the hyperspectral scene, α is a parameter that indicates the position of a given wavelength on the spectral plane, and Λ is the spectral range of interest, typically going from 400nm to 700nm. Notice how the variable s appears in the argument of the coded aperture function, indicating the influence of the mask position into the coding of the information. Given that the sensor is composed of a discrete array of elements, a discretization process of Eq. (2.2) is a crucial step in

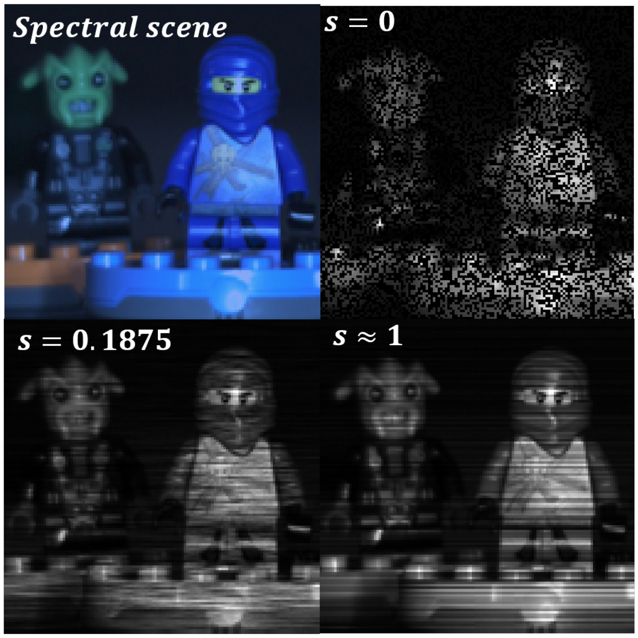

Figure 2.4: Original hyperspectral scene (RGB-mapping), and coded measurements captured at the sensor for 3 different values of s. As s increases, the stretching of the coded aperture pixels is more notorious; when $s = 0$, no stretching is present.

understanding the physical phenomena in terms of the real captured information; this will be done in the following sections of the chapter.

2.3 SSCSI spatial and spectral resolution limits

This section explains in detail the developed analysis to determine the SSCSI spatial and spectral resolution limits. The analysis has been decoupled into two independent problems by assuming a single wavelength scene and a infinitesimally small spatial point, as seen in Fig. 2.5.

2.3.1 Rewriting the continuous model

Consider the $(m, n)^{th}$ pixel detector, where $0 \leq m, n \leq N_d - 1$ are integer and unitless values, and $N_d \times N_d$ are the dimensions of the sensor array. The captured data

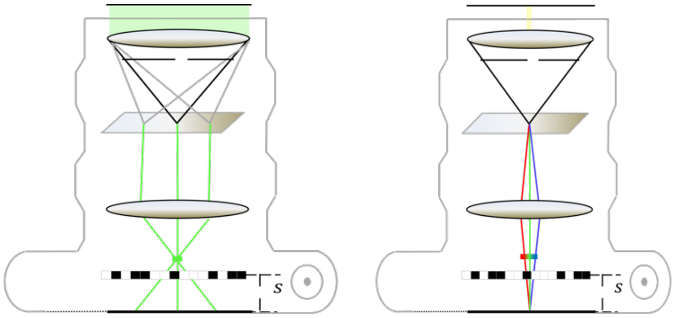

Figure 2.5: Left: single wavelength scene assumed to determine the spatial resolution limits in SSCSI. Right: infinitesimally small spatial point containing all the wavelengths, assumed to determine the spectral resolution limits in SSCSI.

on (m, n) can be written as follows

$$\mathbf{g}_{m,n} = \int_x \int_y \int_\Lambda T(x(1-s)+s\alpha\lambda, y) f(x, y, \lambda) \times \text{rect}\left(\frac{x}{\Delta_d} - m, \frac{y}{\Delta_d} - n\right) d\lambda dy dx, \quad (2.3)$$

where Δ_d is the pitch size of a single detector element and $\text{rect}(x)$ is a rectangular function defined as follows

$$\text{rect}(x, y) = \begin{cases} 1 & \text{If} \quad 0 \leq x, y \leq 1 \\ 0 & \quad \text{otherwise.} \end{cases} \quad (2.4)$$

Likewise, the coded aperture can be written as

$$T(x, y) = \sum_{m'} \sum_{n'} \mathbf{t}_{m',n'} \text{rect}\left(\frac{x}{\Delta_c} - m', \frac{y}{\Delta_c} - n'\right), \quad (2.5)$$

where $\mathbf{t}_{m',n'} \in \{0, 1\} \quad \forall\, (m', n'),\ 0 \leq m', n' \leq N_c - 1$, and $N_c \times N_c$ are the dimensions of the coded aperture mask. The variable Δ_c is the pitch size of the coded aperture. In this dissertation, the coded aperture full width is assumed to be the same as the detector full width, or $N_c\Delta_c = N_d\Delta_d$. Moreover, the detector and the mask are assumed to be totally aligned with respect to each other. Replacing (2.5) into (2.3) leads to

$$\mathbf{g}_{m,n} = \int_x \int_y \int_\Lambda \sum_{m'} \sum_{n'} \mathbf{t}_{m',n'} \times f(x, y, \lambda) \text{rect}\left(\frac{x}{\Delta_d} - m, \frac{y}{\Delta_d} - n\right)$$
$$\times \text{rect}\left(\frac{x(1-s) + s\alpha\lambda}{\Delta_c} - m', \frac{y}{\Delta_c} - n'\right) d\lambda dy dx. \quad (2.6)$$

Notice that the rectangular functions representing the detector and coded aperture elements have dimensions of $\Delta_d \times \Delta_d$ and $\frac{\Delta_c}{1-s} \times \Delta_c$ respectively.

2.3.2 Spatial resolution limits

As mentioned before, to determine the minimum attainable spatial resolution, a single wavelength scene $f(x, y, \lambda_o)\delta(\lambda - \lambda_o)$ is assumed, such that the absence of spectral information simplifies the analysis; here $\delta(\lambda - \lambda_o)$ is a Dirac delta function located at λ_o. By doing that, Eq. (2.6) can be written as follows:

$$\begin{aligned}\mathbf{g}_{m,n} = \int\limits_x \int\limits_y \int\limits_\Lambda \sum_{m'} \sum_{n'} \mathbf{t}_{m',n'} \times f(x, y, \lambda_o) \text{rect}\left(\frac{x}{\Delta_d} - m, \frac{y}{\Delta_d} - n\right) \\ \times \text{ rect}\left(\frac{x(1-s) + s\alpha\lambda}{\Delta_c} - m', \frac{y}{\Delta_c} - n'\right) \delta(\lambda - \lambda_o) d\lambda dy dx. \quad (2.7)\end{aligned}$$

By solving the integral over $d\lambda$, Eq. (2.7) can be rewritten as

$$\begin{aligned}\mathbf{g}_{m,n} = \int\limits_x \int\limits_y \sum_{m'} \sum_{n'} \mathbf{t}_{m',n'} \times f(x, y, \lambda_o) \text{rect}\left(\frac{x}{\Delta_d} - m, \frac{y}{\Delta_d} - n\right) \\ \times \text{ rect}\left(\frac{x(1-s) + s\alpha\lambda_o}{\Delta_c} - m', \frac{y}{\Delta_c} - n'\right) dy dx. \quad (2.8)\end{aligned}$$

The minimum spatially resolvable feature, given by Eq. (2.8) can be determined by establishing the intersecting region of the coded aperture and detector rectangular functions, which is going to depend on the mask pixel stretching, as explained below.

2.3.2.1 Stretched coded aperture pitch size smaller than detector pitch size: $\frac{\Delta_c}{1-s} \leq \Delta_d$

In the first scenario, it is assumed that the stretched coded aperture pixel length does not exceed the detector pitch size. Figure 2.6-left shows the projection of the pattern onto the sensor for this case. By using the same principle proposed for CASSI [30, 31], the minimum attainable spatial resolution is defined by $\frac{\Delta_c}{1-s} \times \Delta_c$, which implies that several spatial pixels of the scene will be registered by a single detector pixel; this allows the recovery of spatially super-resolved datacubes. To model the sensing process, the overlapping region between detector and coded aperture pixels must be defined. Let (m'_l, n'_l) and (m'_r, n'_r) be the upper-leftmost and lower-rightmost coded aperture elements impinging on the $(m, n)^{th}$ detector element. The overlapping of the two rectangular functions in Eq. (2.8) can be seen in Figs. 2.7 and 2.8. From Fig. 2.7, it can be inferred that for the overlap to exist on the y axis, the following condition must hold

$$n\frac{\Delta_d}{\Delta_c} \leq n' \leq (n+1)\frac{\Delta_d}{\Delta_c} - 1, \tag{2.9}$$

where the limits in the last inequality come from solving $n'_l\Delta_c = n\Delta_d$ and $(n'_r+1)\Delta_c = (n+1)\Delta_d$ for n'_l and n'_r respectively. Likewise, it can be shown based on Fig. 2.8, that for the two rectangular functions in Eq. (2.8) to overlap in the x axis, the next condition must hold

$$\left\lfloor \frac{(m)\Delta_d(1-s) + s\alpha\lambda_o}{\Delta_c} \right\rfloor \leq m' \leq \left\lfloor \frac{(m+1)\Delta_d(1-s) + s\alpha\lambda_o}{\Delta_c} \right\rfloor, \tag{2.10}$$

where $\lfloor \cdot \rfloor$ is the floor operator. The limits in the last inequality come from solving $\frac{m'_l\Delta_c - s\alpha\lambda_o}{1-s} \leq m\Delta_d \leq \frac{(m'_l+1)\Delta_c - s\alpha\lambda_o}{1-s}$ and $\frac{m'_r\Delta_c - s\alpha\lambda_o}{1-s} \leq (m+1)\Delta_d \leq \frac{(m'_r+1)\Delta_c - s\alpha\lambda_o}{1-s}$ for m'_l and m'_r respectively and considering that m'_l and m'_r are integer indexes. Given that the minimum resolvable spatial feature is of dimensions $\frac{\Delta_c}{(1-s)} \times \Delta_c$, the captured information in the $(m, n)^{th}$ detector element, $\mathbf{g}_{m,n}$, can be represented as

$$\mathbf{g}_{m,n} = \sum_{m'=m'_l}^{m'_r} \sum_{n'=n\frac{\Delta_d}{\Delta_c}}^{(n+1)\frac{\Delta_d}{\Delta_c}-1} W_{m,m'} \times \mathbf{t}_{m',n'} \times \mathbf{f}_{m',n',\lambda_o}, \tag{2.11}$$

where m'_l and m'_r are the limits of the expression in Eq. (2.10) and the variable $\mathbf{f}$ is the datacube representation of the hyperspectral scene, being $\mathbf{f}_{m',n',\lambda_o}$ the information of the datacube at position (m', n') and spectral band λ_o; $\mathbf{f}_{m',n',\lambda_o}$ is obtained by applying the integral operators in Eq. (2.8). Notice that the spatial dimensions of $\mathbf{f}$, using a $N_d \times N_d$ sensor array, are defined as $\left\lceil N_d\frac{\Delta_d}{\Delta_c/(1-s)} \right\rceil \times N_d\frac{\Delta_d}{\Delta_c}$, where $\lceil \cdot \rceil$ is the ceiling operator. The parameter $W_{m,m'}$ is the fraction of the coded aperture element $\mathbf{t}_{m',n'}$

impinging on the $(m,n)^{th}$ detector element, and it is defined as follows (see appendix A for full derivation),

$$W_{m,m'} = \begin{cases} \frac{\frac{(m'+1)\Delta_c - s\alpha\lambda_o}{1-s} - (m)\Delta_d}{\Delta_c/(1-s)} & \text{If} \quad m' = m'_l \\ \frac{(m+1)\Delta_d - \frac{m'\Delta_c - s\alpha\lambda_o}{1-s}}{\Delta_c/(1-s)} & \text{If} \quad m' = m'_r \\ 1 & \text{If} \quad m'_l < m' < m'_r \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

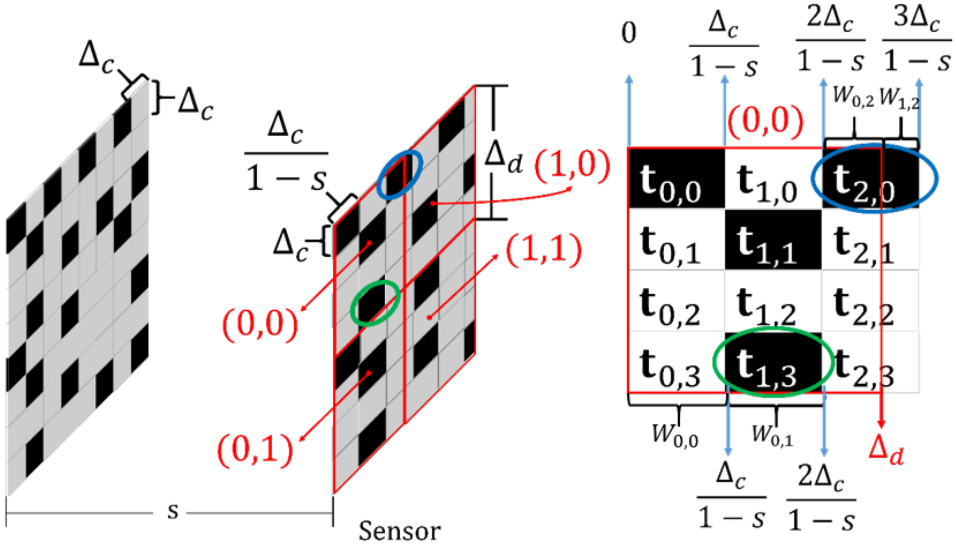

Figure 2.6: Left: Original Coded aperture at a given position s and projection of the coded aperture onto the sensor. Here $\frac{\Delta_c}{1-s} \leq \Delta_d$. The red squares represent the sensor elements $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$. As depicted, a coded aperture pixel of size $\Delta_c \times \Delta_c$ is mapped into a $\frac{\Delta_c}{1-s} \times \Delta_c$ rectangle on the detector. Right: Front view of the sensor element (0,0) and illustration of the W parameters of Eq. (2.11). The W parameters represent the percentage of a coded aperture pixel impinging on a given detector element. Notice, for example, that one of the portions of $\mathbf{t}_{2,0}$ impinges on the $(1,0)^{th}$ sensor element; this portion is represented by $W_{1,2}$ (see appendix A for full explanation).

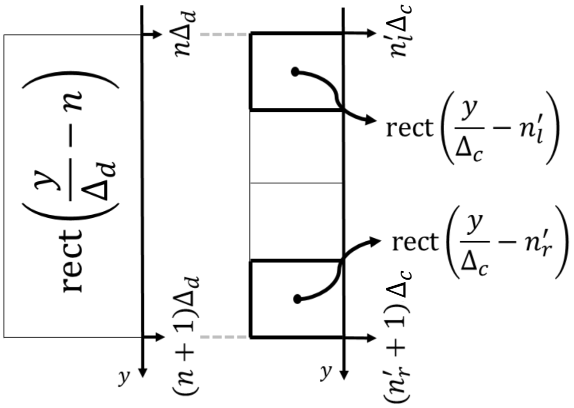

Figure 2.7: Overlapping area in the y axis of the two rectangular functions given in Eq. (2.8), for a given n. Here, it is assumed that the ratio Δ_d/Δ_c is an integer greater than or equal to 1. Notice that $n_l'\Delta_c = n\Delta_d$ and $(n_r'+1)\Delta_c = (n+1)\Delta_d$.

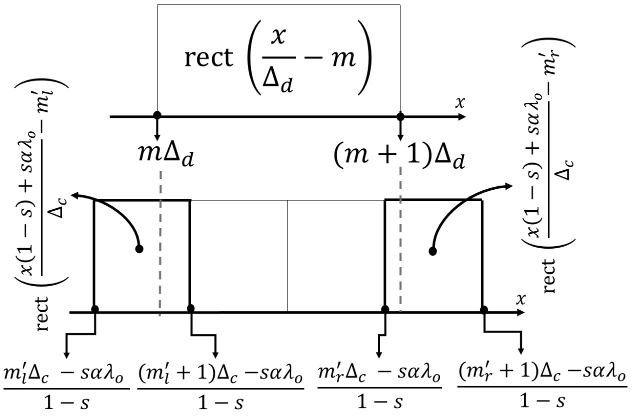

Figure 2.8: Overlapping area in the x axis of the two rectangular functions in Eq. (2.8), for a given m. Notice that $\frac{m_l'\Delta_c - s\alpha\lambda_o}{1-s} \leq m\Delta_d \leq \frac{(m_l'+1)\Delta_c - s\alpha\lambda_o}{1-s}$ and $\frac{m_r'\Delta_c - s\alpha\lambda_o}{1-s} \leq (m+1)\Delta_d \leq \frac{(m_r'+1)\Delta_c - s\alpha\lambda_o}{1-s}$.

2.3.2.2 Stretched coded aperture pitch size greater than detector pitch size: $\frac{\Delta_c}{1-s} > \Delta_d$

The projection of the coded aperture into the detector for this scenario can be seen in Fig. 2.9-left. Unlike the previous scenario, here the spatial resolution is limited by the detector pitch size and defined as $\Delta_d \times \Delta_c$. Again, a coded aperture-detector mismatch is present, and in order to model this particular phenomenon, a synthetic gray-scale pattern is created through the combination of the pixels affecting the same region, as illustrated by the green circle in Fig. 2.9. Given the $(m', n')^{th}$ coded aperture pixel and the $(m, n)^{th}$ detector element, the two rectangular functions in Eq. (2.8) intersect each other if n' falls within the interval specified in Eq. (2.9) and $m' = \left\lfloor \frac{(m)\Delta_d(1-s)+s\alpha\lambda_o}{\Delta_c} \right\rfloor + 1$ (see appendix B). Taking into account that the minimum spatially resolvable feature is of dimensions $\Delta_d \times \Delta_c$, the captured information in a particular sensor element, $\mathbf{g}_{m,n}$, can be written as follows

$$\mathbf{g}_{m,n} = \sum_{n'=n\frac{\Delta_d}{\Delta_c}}^{(n+1)\frac{\Delta_d}{\Delta_c}-1} (\mathbf{t}_{m'-1,n'} \times p_m + \mathbf{t}_{m',n'} \times (1-p_m)) \times \mathbf{f}_{m,n',\lambda_o}, \tag{2.13}$$

where $\mathbf{f}_{m,n',\lambda_o}$ is obtained by applying the integral operators in Eq. (2.8). Notice that the spatial dimensions of $\mathbf{f}$, using an $N_d \times N_d$ sensor array, are defined as $N_d \times N_d \frac{\Delta_d}{\Delta_c}$. The expression $\tilde{\mathbf{t}}_{m,n'} = \mathbf{t}_{m'-1,n'} \times p_m + \mathbf{t}_{m',n'} \times (1-p_m)$ represents the effective coded aperture, where p_m indicates the percentage occupied by $\mathbf{t}_{m'-1,n'}$ on the $(m, n)^{th}$ detector element and is defined as follows (see appendix B for full derivation),

$$p_m = \begin{cases} 1 & \text{If} \quad (m+1)\Delta_d \leq \frac{m'\Delta_c - s\alpha\lambda_o}{(1-s)} \\ \frac{\frac{m'\Delta_c - s\alpha\lambda_o}{(1-s)} - (m)\Delta_d}{\Delta_d} & \text{If} \quad (m+1)\Delta_d > \frac{m'\Delta_c - s\alpha\lambda_o}{(1-s)}. \end{cases} \tag{2.14}$$

With the spatial resolution limits defined, it is now proceeded to calculate the SSCSI spectral resolution limits.

2.3.3 Spectral resolution limits

As previously mentioned, the SSCSI spectral resolution must be calculated assuming an infinitesimally small spatial point containing all the wavelengths $f(x, y, \lambda)\delta(x-$

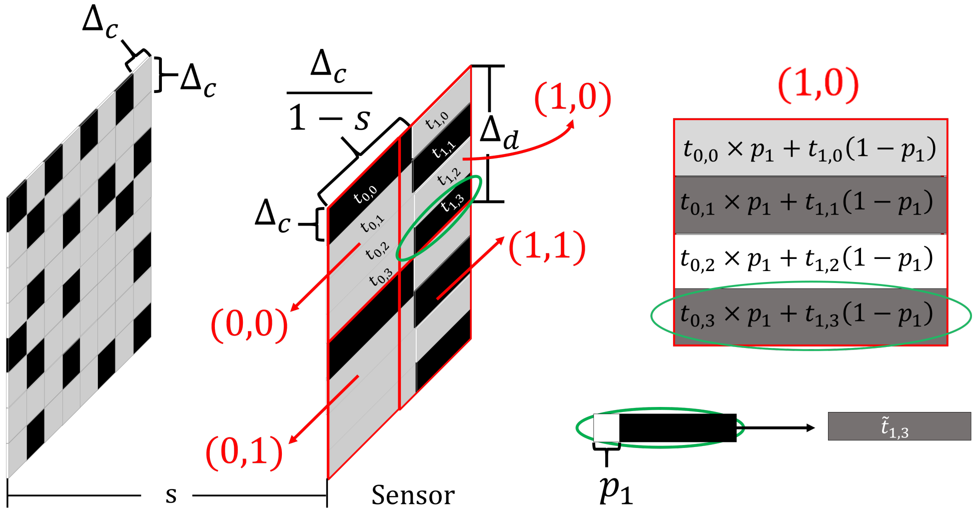

Figure 2.9: Left: Coded aperture at a given position s and projection of the coded aperture onto the sensor. Here $\frac{\Delta_c}{1-s} > \Delta_d$. The red squares represent the sensor elements $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$. As depicted, a coded aperture pixel of size $\Delta_c \times \Delta_c$ is mapped into a $\frac{\Delta_c}{1-s} \times \Delta_c$ rectangle on the detector. Right: Effective coded aperture $\hat{\mathbf{t}}$ impinging in the $(1,0)^{th}$ sensor element. Notice, in the green circled region, that the variable p_1 indicates the area of the sensor occupied by $\mathbf{t}_{0,3}$.

$x_o, y-y_o)$, where $\delta(x-x_o, y-y_o)$ represents a Dirac function located at the point (x_o, y_o). With this assumption, Equation (2.6) can be rewritten as follows

$$\mathbf{g}_{m,n} = \int_x \int_y \int_\Lambda \sum_{m'} \sum_{n'} \mathbf{t}_{m',n'} \times f(x_o, y_o, \lambda) \mathrm{rect}\left(\frac{x}{\Delta_d} - m, \frac{y}{\Delta_d} - n\right) \times \mathrm{rect}\left(\frac{x(1-s)+s\alpha\lambda}{\Delta_c} - m', \frac{y}{\Delta_c} - n'\right) \times \delta(x - x_o, y - y_o) d\lambda dy dx. \quad (2.15)$$

By solving the integral over $dxdy$, Equation (2.15) can be rewritten as

$$\mathbf{g}_{m,n} = \int_\Lambda \sum_{m'} \sum_{n'} \mathbf{t}_{m',n'} \times f(x_o, y_o, \lambda) \mathrm{rect}\left(\frac{x_o}{\Delta_d} - m, \frac{y_o}{\Delta_d} - n\right) \times \mathrm{rect}\left(\frac{x_o(1-s)+s\alpha\lambda}{\Delta_c} - m', \frac{y_o}{\Delta_c} - n'\right) d\lambda. \quad (2.16)$$

The attainable spectral resolution is found by defining the region where the two rectangular functions in the last equation intersect each other. This ensures the integral will not equal zero. The analysis was again divided in two cases as elaborated next.

2.3.3.1 Stretched coded aperture pitch size smaller than detector pitch size: $\frac{\Delta_c}{1-s} \leq \Delta_d$

Figure 2.10 shows the overlapping of the two rectangular functions in Eq. (2.16), when $\frac{\Delta_c}{1-s} \leq \Delta_d$; three different regions can be noted. When the overlapping region is the gray area, $m\Delta_d \leq \frac{(m'_1+1)\Delta_c - s\alpha\lambda}{1-s}$ and $m\Delta_d \geq \frac{m'_1\Delta_c - s\alpha\lambda}{1-s}$ or, in other words, $\frac{m'_1\Delta_c - m\Delta_d(1-s)}{s\alpha} \leq \lambda \leq \frac{(m'_1+1)\Delta_c - m\Delta_d(1-s)}{s\alpha}$. This interval has an extension of $\Delta_\lambda = \frac{\Delta_c}{s\alpha}$. When the overlapping region is the yellow one, $(m+1)\Delta_d \leq \frac{(m'_2+2)\Delta_c - s\alpha\lambda}{1-s}$ and $(m+1)\Delta_d \geq \frac{(m'_2+1)\Delta_c - s\alpha\lambda}{1-s}$ or, in other words, $\frac{(m'_2+1)\Delta_c - (m+1)\Delta_d(1-s)}{s\alpha} \leq \lambda \leq \frac{(m'_2+2)\Delta_c - (m+1)\Delta_d(1-s)}{s\alpha}$. This interval has also an extension of $\Delta_\lambda = \frac{\Delta_c}{s\alpha}$. Likewise, when the overlapping region is the green area, $m\Delta_d \leq \frac{m'_2\Delta_c - s\alpha\lambda}{1-s}$ and $(m+1)\Delta_d \geq \frac{(m'_2+1)\Delta_c - s\alpha\lambda}{1-s}$, or equivalently $\frac{(m'_2+1)\Delta_c - (m+1)\Delta_d(1-s)}{s\alpha} \leq \lambda \leq \frac{m'_2\Delta_c - m\Delta_d(1-s)}{s\alpha}$, which has an extension of $\Delta_\lambda = \frac{\Delta_d(1-s)-\Delta_c}{s\alpha}$. Let the ratio $\Delta_d/\Delta_c = C$, where C is an integer greater than 1. In order to make $\frac{\Delta_c}{s\alpha} > \frac{\Delta_d(1-s)-\Delta_c}{s\alpha}$, one can prove that $s > 1 - \frac{2}{C}$. For $C = 2$, the inequality $\frac{\Delta_c}{s\alpha} > \frac{\Delta_d(1-s)-\Delta_c}{s\alpha}$ holds for $s > 0$. Therefore, $\Delta_\lambda = \frac{\Delta_c}{s\alpha}$ can be taken as an upper bound of the spectral resolution. When $C > 2$, $\frac{\Delta_c}{s\alpha} < \frac{\Delta_d(1-s)-\Delta_c}{s\alpha}$ for certain values of s, and $\Delta_\lambda = \frac{\Delta_c}{s\alpha}$ represents the exact spectral resolution given by the SSCSI.

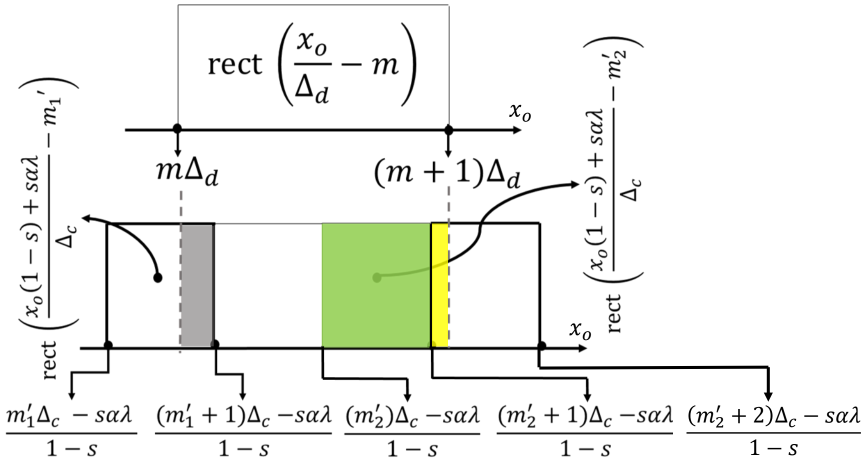

Figure 2.10: Graphical explanation to calculate the spectral resolution of the SSCSI, when $\frac{\Delta_c}{1-s} \leq \Delta_d$.

2.3.3.2 Stretched coded aperture pitch size greater than detector pitch size: $\frac{\Delta_c}{1-s} > \Delta_d$

Figure 2.11 shows the overlapping of the two rectangular functions, when $\frac{\Delta_c}{1-s} > \Delta_d$; two cases must be analyzed. In the first one, the gray area partially occupies the sensor element, which implies that $m\Delta_d \leq \frac{(m'+1)\Delta_c - s\alpha\lambda}{1-s}$ and $(m+1)\Delta_d \geq \frac{(m'+1)\Delta_c - s\alpha\lambda}{1-s}$ or, in other words, $\frac{(m'+1)\Delta_c-(m+1)\Delta_d(1-s)}{s\alpha} \leq \lambda \leq \frac{(m'+1)\Delta_c-(m)\Delta_d(1-s)}{s\alpha}$, whose extension is equal to $\Delta_\lambda = \frac{\Delta_d(1-s)}{s\alpha}$. In the second case, the gray area fully occupies the sensor element, which means that $m\Delta_d \geq \frac{(m')\Delta_c - s\alpha\lambda}{1-s}$ and $(m+1)\Delta_d \leq \frac{(m'+1)\Delta_c - s\alpha\lambda}{1-s}$, or equivalently $\frac{(m')\Delta_c-(m)\Delta_d(1-s)}{s\alpha} \leq \lambda \leq \frac{(m'+1)\Delta_c-(m+1)\Delta_d(1-s)}{s\alpha}$; this interval has an extension of $\Delta_\lambda = \frac{\Delta_c-\Delta_d(1-s)}{s\alpha}$. Notice that, since $\frac{\Delta_c}{1-s} > \Delta_d$, $\frac{\Delta_c}{s\alpha} > \frac{\Delta_d(1-s)}{s\alpha}$ and $\frac{\Delta_c}{s\alpha} > \frac{\Delta_c-\Delta_d(1-s)}{s\alpha}$. Therefore, $\frac{\Delta_c}{s\alpha}$ can be taken as an upper bound of the spectral resolution.

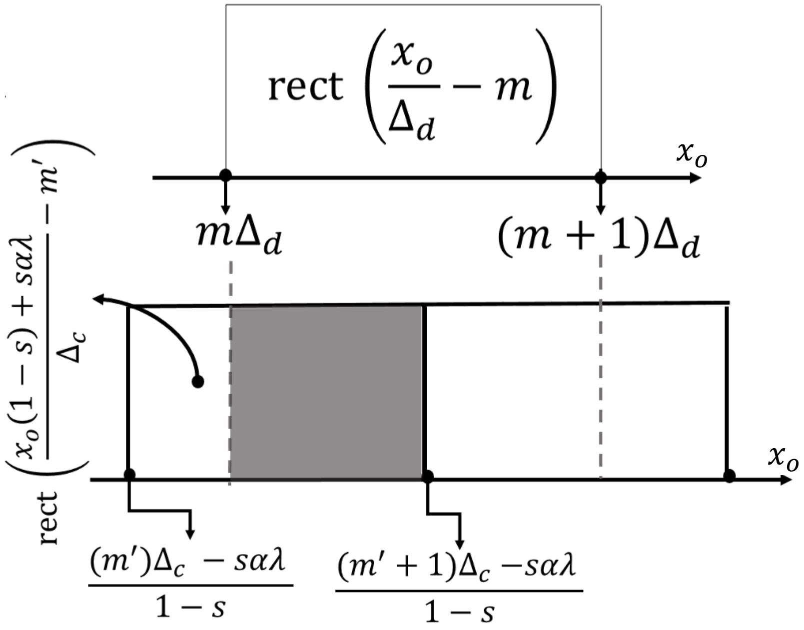

Figure 2.11: Graphical explanation to calculate the spectral resolution of the SSCSI, when $\frac{\Delta_c}{1-s} > \Delta_d$.

2.3.3.3 Number of resolvable spectral bands

With $\Delta_\lambda = \frac{\Delta_c}{s\alpha}$ defined as an upper bound for the spectral resolution that encompasses all the scenarios, the number of resolvable spectral bands can be defined as follows,

$$L = \left\lceil \frac{s\alpha\left(\lambda_{max} - \lambda_{min}\right)}{\Delta_c} \right\rceil, \tag{2.17}$$

where $\Lambda = \lambda_{max} - \lambda_{min}$ is the spectral range of interest. The term $s\alpha\left(\lambda_{max} - \lambda_{min}\right)$ indicates the spectral dispersion of the scene at a given position s. When the coded aperture and the detector width are the same ($N_c\Delta_c = N_d\Delta_d$), the last expression can be rewritten as follows

$$L = \left\lceil s\frac{N_d}{\Delta_c/\Delta_d}\beta \right\rceil, \tag{2.18}$$

where $\beta = \alpha(\lambda_{max} - \lambda_{min})/N_c\Delta_c$ is the ratio between the spectral plane length and the coded aperture width.

2.3.3.4 Coding process between adjacent bands

Let $\lambda_k = \lambda_{min} + k\frac{\Delta_c}{s\alpha}$ and $\lambda_{k+1} = \lambda_{min} + (k+1)\frac{\Delta_c}{s\alpha}$ be two resolvable spectral bands such that $\lambda_{k+1} - \lambda_k = \frac{\Delta_c}{s\alpha}$. The physical separation between these bands when arriving at the coded aperture plane is equal to $s\alpha\left(\lambda_{k+1} - \lambda_k\right) = \Delta_c$, which means that the two adjacent bands will be coded by the same coded aperture, but shifted by one column. A graphical explanation of the coding process in adjacent bands can be seen in Fig. 2.12.

2.3.4 Fusing the spatial and spectral resolution analysis

The spectral resolution limits must now be incorporated into the discrete measurements models given by Eqs. (2.11) and (2.13) and take into consideration the shifted-by-one-column coding process between adjacent bands. The final discrete measurements model when $\frac{\Delta_c}{1-s} \leq \Delta_d$ can be rewritten as follows:

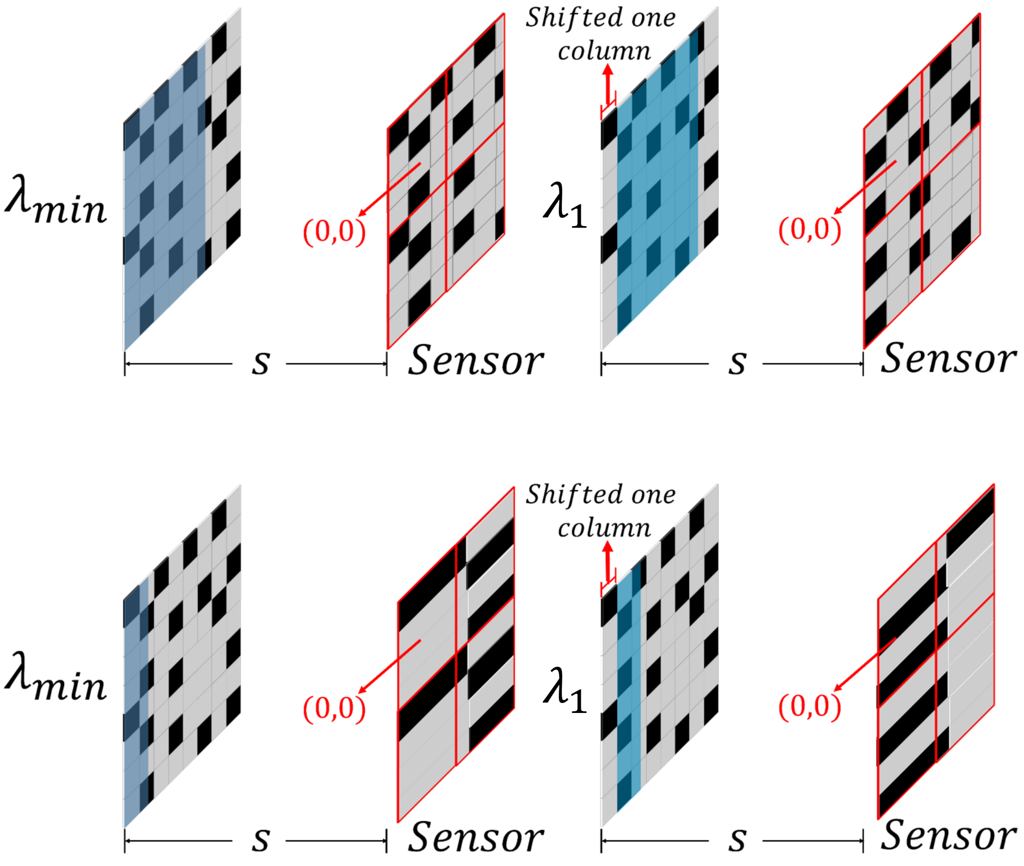

Figure 2.12: Top: Coding process, at a given s, for two adjacent and resolvable bands λ_{min} and $\lambda_1 = \lambda_{min} + \frac{\Delta_c}{s\alpha}$, when $\Delta_c/(1-s) \leq \Delta_d$. Bottom: Coding process, at a given s, for two adjacent and resolvable bands λ_{min} and $\lambda_1 = \lambda_{min} + \frac{\Delta_c}{s\alpha}$, when $\Delta_c/(1-s) > \Delta_d$. Notice that, in both cases λ_{min} impinges on the leftmost side of the coded aperture. The dispersion between λ_1 and λ_{min} is equal to $s\alpha(\lambda_1 - \lambda_{min}) = \Delta_c$. This means they are coded by the same coded aperture but shifted by one column.

$$\mathbf{g}_{m,n} = \sum_{m'=m'_l}^{m'_r} \sum_{n'=n\frac{\Delta_d}{\Delta_c}}^{(n+1)\frac{\Delta_d}{\Delta_c}-1} \sum_{k=0}^{L-1} W_{m,m'} \mathbf{t}_{m'+k,n'} \mathbf{f}_{m',n',k}, \tag{2.19}$$

where $m'_l = \left\lfloor \frac{(m)\Delta_d(1-s)+s\alpha\lambda_{min}}{\Delta_c} \right\rfloor$, $m'_r = \left\lfloor \frac{(m+1)\Delta_d(1-s)+s\alpha\lambda_{min}}{\Delta_c} \right\rfloor$ and $W_{m,m'}$ is defined in Eq. (2.12) (for $\lambda_o = \lambda_{min}$). The k^{th} spectral band is defined here as the interval $[\lambda_{min} + \frac{\Delta_c}{s\alpha}k, \lambda_{min} + \frac{\Delta_c}{s\alpha}(k+1)]$; the shifting by one column between adjacent bands can be seen in the “ $+\,k$” term on the coded aperture array $\mathbf{t}$. The dimensions of the recovered datacube are given by

$$\left\lceil N_d \frac{\Delta_d}{\Delta_c/(1-s)} \right\rceil \times N_d \frac{\Delta_d}{\Delta_c} \times L, \tag{2.20}$$

where L is given in Eqs. (2.17) and (2.18). As an example, take the sensor element $(0,0)$ in Fig. 2.12-top; the captured information in that element related to the spectral band $[\lambda_{min}, \lambda_{min} + \frac{\Delta_c}{s\alpha}]$ will be given by

$$\mathbf{g}_{0,0} = \sum_{m'=0}^{2} \sum_{n'=0}^{3} W_{0,m'} \mathbf{t}_{m',n'} \mathbf{f}_{m',n',0}. \tag{2.21}$$

Likewise, the captured information for the spectral band $[\lambda_{min} + \frac{\Delta_c}{s\alpha}, \lambda_{min} + \frac{2\Delta_c}{s\alpha}]$ will be given by

$$\mathbf{g}_{0,0} = \sum_{m'=0}^{2} \sum_{n'=0}^{3} W_{0,m'} \mathbf{t}_{m'+1,n'} \mathbf{f}_{m',n',1}. \tag{2.22}$$

In this particular case, $W_{0,0} = 1$ and $W_{0,1} = 1$. On the other side, the final discrete measurements model when $\frac{\Delta_c}{1-s} > \Delta_d$ can be seen below:

$$\mathbf{g}_{m,n} = \sum_{n'=n\frac{\Delta_d}{\Delta_c}}^{(n+1)\frac{\Delta_d}{\Delta_c}-1} \sum_{k=0}^{L-1} \left(\mathbf{t}_{m'+k-1,n'} \times p_m + \mathbf{t}_{m'+k,n'} \times (1-p_m)\right) \mathbf{f}_{m,n',k}, \tag{2.23}$$

where the k^{th} spectral band is defined as the interval $[\lambda_{min} + \frac{\Delta_c}{s\alpha}k, \lambda_{min} + \frac{\Delta_c}{s\alpha}(k+1)]$, $m' = \left\lfloor \frac{(m)\Delta_d(1-s)+s\alpha\lambda_{min}}{\Delta_c} \right\rfloor + 1$ and p_m is as defined in Eq. (2.14) (for $\lambda_o = \lambda_{min}$). The shifting by one column between adjacent bands can be seen in the “ $+\,k$” term on the coded aperture array $\mathbf{t}$. The dimensions of the recovered datacube are given by

$$N_d \times N_d \frac{\Delta_d}{\Delta_c} \times L, \tag{2.24}$$

where L is given in Eqs. (2.17) and (2.18). As an example, take the sensor element $(1,0)$ in Fig. 2.12-bottom. The captured information in that element related to the spectral band $[\lambda_{min}, \lambda_{min} + \frac{\Delta_c}{s\alpha}]$ will be given by

$$\mathbf{g}_{1,0} = \sum_{n'=0}^{3} \left(\mathbf{t}_{0,n'} \times p_1 + \mathbf{t}_{1,n'} \times (1-p_1)\right) \mathbf{f}_{1,n',0}. \tag{2.25}$$

Likewise, the captured information for the spectral band $[\lambda_{min} + \frac{\Delta_c}{s\alpha}, \lambda_{min} + \frac{2\Delta_c}{s\alpha}]$ will be given by

$$\mathbf{g}_{1,0} = \sum_{n'=0}^{3} \left(\mathbf{t}_{1,n'} \times p_1 + \mathbf{t}_{2,n'} \times (1-p_1)\right) \mathbf{f}_{1,n',1}. \tag{2.26}$$

2.3.5 Summary

The summary of the spatial and spectral resolution limits and the discrete measurements models are provided in Tables 2.1, 2.2 and 2.3.

Table 2.1: SSCSI spatial resolution limits.

	$\frac{\Delta_c}{(1-s)} \leq \Delta_d$	$\frac{\Delta_c}{(1-s)} > \Delta_d$
Spatial resolution	$\frac{\Delta_c}{(1-s)} \times \Delta_c$	$\Delta_d \times \Delta_c$

Table 2.2: SSCSI spectral resolution limits. An upper bound of $\Delta_\lambda = \frac{\Delta_c}{s\alpha}$ that encompasses all the cases was defined to simplify the analysis.

	$\frac{\Delta_c}{(1-s)} \leq \Delta_d$	$\frac{\Delta_c}{(1-s)} > \Delta_d$
Spectral resolution	$\frac{\Delta_c}{s\alpha}, \frac{\Delta_d(1-s)-\Delta_c}{s\alpha}$	$\frac{\Delta_d(1-s)}{s\alpha}, \frac{\Delta_c-\Delta_d(1-s)}{s\alpha}$

Table 2.3: Proposed discrete measurements models.

$\frac{\Delta_c}{(1-s)} \leq \Delta_d$	$\mathbf{g}_{m,n} = \sum_{m'=m'_l}^{m'_r} \sum_{n'=n\frac{\Delta_d}{\Delta_c}}^{(n+1)\frac{\Delta_d}{\Delta_c}-1} \sum_{k=0}^{L-1} W_{m,m'} \mathbf{t}_{m'+k,n'} \mathbf{f}_{m',n',k}$
$\frac{\Delta_c}{(1-s)} > \Delta_d$	$\mathbf{g}_{m,n} = \sum_{n'=n\frac{\Delta_d}{\Delta_c}}^{(n+1)\frac{\Delta_d}{\Delta_c}-1} \sum_{k=0}^{L-1} \left(\mathbf{t}_{m'+k-1,n'} \times p_m + \mathbf{t}_{m'+k,n'} \times (1-p_m)\right) \mathbf{f}_{m,n',k}$

2.3.6 Exact coded aperture shearing

The proposed models are an approximation which do not account for the exact projection of the coded aperture into the sensor. To understand this in detail, one must recall Eq. (2.6), which characterizes the SSCSI sensing process. By analyzing the x dimension of the coded aperture rectangular function in this expression, given by $\frac{x(1-s)+s\alpha\lambda}{\Delta_c} - m'$, one can obtain Fig. 2.13-top, that depicts the actual sheared pattern in the $x - \lambda$ plane. The dashed black lines in the figure represent the limits of the m^{th} detector, while the dashed red lines are the spectral divisions every $\Delta_\lambda = \frac{\Delta_c}{s\alpha}$. The connection between the exact shearing and the proposed models in Eqs. (2.19) and (2.23) can be seen in Fig. 2.13-bottom. As depicted, the spectral range $(\lambda_1 - \lambda_2)$ is uniformly coded as λ_1. Hence, the proposed discretization model must be considered a *first order approximation*.

2.4 Simulations

2.4.1 Matrix forward model

The SSCSI sensing process can be described using matrix-vector multiplication notation, as follows:

$$\mathbf{g}^{(q)} = \mathbf{H}^{(q)}\boldsymbol{\Psi}\boldsymbol{\pi} \quad q = 1, 2, ..., Q, \tag{2.27}$$

where Q is the number of captured shots and $\mathbf{g}^{(q)}$ are the vectorized measurements of length N_d^2. The variable $\boldsymbol{\pi}$ is a sparse vector representation of the datacube $\mathbf{f}$ in a basis $\boldsymbol{\Psi}$. If the datacube to be recovered has dimensions $N_x \times N_y \times L$, then the length of $\boldsymbol{\pi}$ is $N_x N_y L$. The diagonal-structured matrix $\mathbf{H}^{(q)}$ performs the coding process of the datacube; its dimensions are $N_d^2 \times N_x N_y L$. For a multishot approach (several captured snapshots with different patterns), the sensing process can be written as

$$\mathbf{g} = \mathbf{H}\boldsymbol{\Psi}\boldsymbol{\pi}, \tag{2.28}$$

where $\mathbf{H} = \left[\mathbf{H}^{(1)^T}, \mathbf{H}^{(2)^T}, \ldots, \mathbf{H}^{(Q)^T}\right]^T$ is a diagonal-structured matrix with dimensions $QN_d^2 \times N_x N_y L$, and $\mathbf{g} = \left[\mathbf{g}^{(1)T}, \mathbf{g}^{(2)T}, \ldots, \mathbf{g}^{(Q)T}\right]^T$ is a vector of length QN_d^2. The

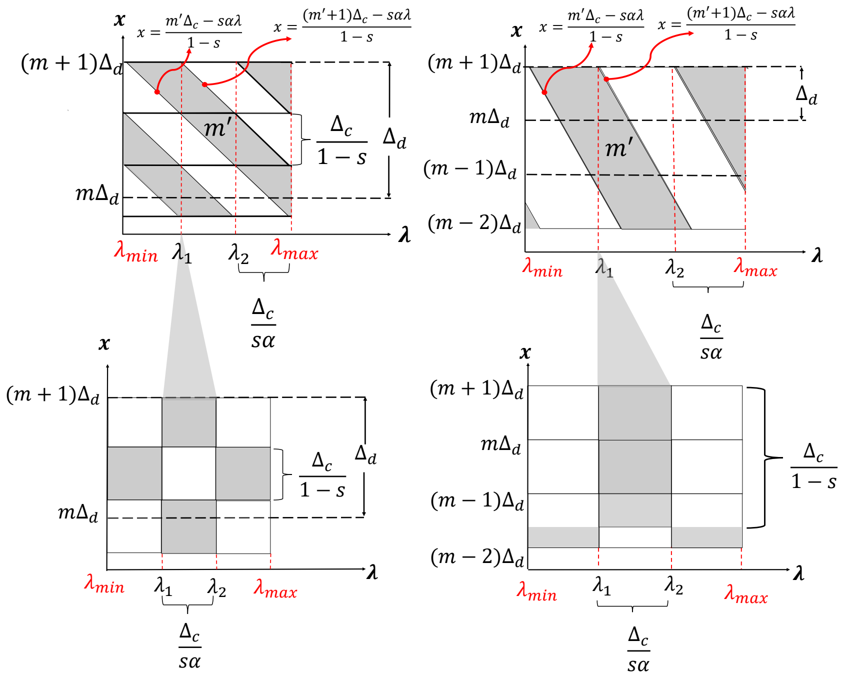

Figure 2.13: Top left: Original sheared coded aperture when $\Delta_c/(1-s) \leq \Delta_d$. Top right: Original sheared coded aperture when $\Delta_c/(1-s) > \Delta_d$. Bottom left: First order approximation when $\Delta_c/(1-s) \leq \Delta_d$. Bottom right: First order approximation when $\Delta_c/(1-s) > \Delta_d$. Notice how the spectral band $[\lambda_1, \lambda_2]$ is coded by the same pattern that codes the hyperspectral scene at λ_1.

structure of the matrix $\mathbf{H}$ for a single shot and a single wavelength, for $\Delta_c/(1-s) \leq \Delta_d$ and $\Delta_c/(1-s) > \Delta_d$ can be seen in Fig. 2.14-left. The structure of $\mathbf{H}$ for $Q = 2$ and $L = 2$ can be seen in Fig. 2.14-right.

2.4.2 Simulation conditions

2.4.2.1 SSCSI detector and coded aperture dimensions

The recovery of a hyperspectral image composed of 24 bands is simulated assuming a detector size of $N_d \times N_d$, where $N_d = 256$. The coded aperture and detector pitch sizes are assumed to be the same; given that $\Delta_c = \Delta_d$, $\Delta_c/(1-s) \geq \Delta_d$ for any value of s; Eq. (2.23) therefore applies. Assuming that $\beta = 1$, which means that the coded aperture width and spectral plane length are the same, $s \approx 0.1$ in order to

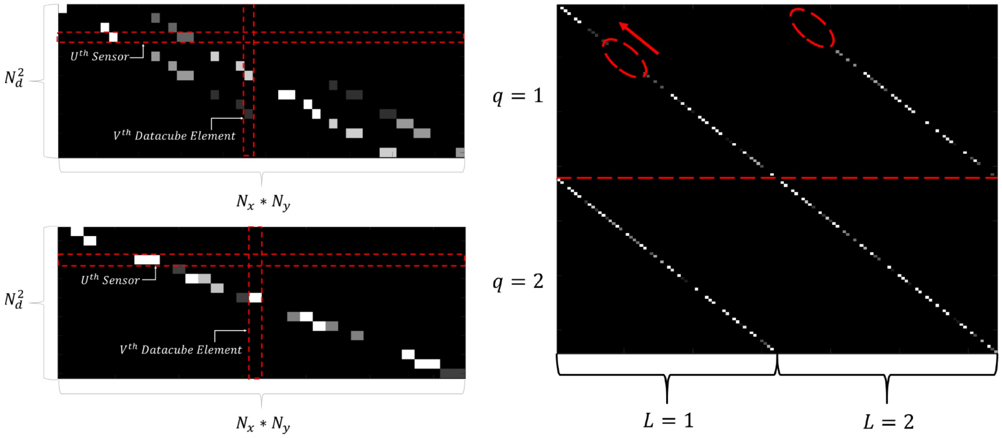

Figure 2.14: Left: Sensing matrix $\mathbf{H}^{(q)}$, for a single shot and a $N_x \times N_y$ single wavelength datacube. Top-left: when $\frac{\Delta_c}{1-s} \leq \Delta_d$. Bottom-left: when $\frac{\Delta_c}{1-s} > \Delta_d$. Each row represents a sensor element and each column represents a datacube element. Right: Sensing matrix for two shots and a two wavelengths datacube, where $\Delta_c = \Delta_d$, $s = 0.18$ and $\beta = 1$. The dashed red horizontal line separates the two different shots, $q = 1$ and $q = 2$. The size of the implemented sensor and coded aperture is 8×8. The dashed red circles and red arrow indicate the dispersion process between bands.

recover the 24 bands (Eq. (2.18)). The final dimensions of the recovered datacube, according to Eq. (2.24), will be equal to $256 \times 256 \times 24$. The coded aperture patterns were generated such that complementarity among shots is preserved, where the complementarity is defined as

$$\sum_{q=1}^{Q} \mathbf{t}^{q}_{m,n} = 1 \; \forall \; m, n. \tag{2.29}$$

To avoid trivial patterns (all-ones, all-zeros), the transmittance of every pattern is inversely proportional to the number of capture snapshots, where the transmittance is defined as the ratio between the number of *ones* and the total number of pixels in the pattern. This structure is called Boolean, and it has been proven to enhance the quality of the reconstructions [32].

2.4.2.2 CASSI and Colored-CASSI

Two state-of-the-art architectures were simulated to test the SSCSI performance, the CASSI and the Colored-CASSI. The CASSI coding patterns were generated using the Boolean structure defined in Eq. (2.29). The colored CASSI was simulated using two different coding principles; in the first one, no limit on the number of optical filters was imposed; the complementary condition for this case is defined as:

$$\sum_{q=1}^{Q} \mathbf{t}^{q}_{m,n,k} = 1 \ \forall \ m, n, k. \tag{2.30}$$

This imager is called *Ideal Colored-CASSI*. The second Colored-CASSI coding principle was simulated using four different predefined optical filters: Low-pass, High-pass, Band-pass, and Band-stop. The optical filters were chosen such that the following condition holds

$$\sum_{q=1}^{Q} \mathbf{t}^{q}_{m,n,k} \geq 1 \ \forall \ m, n, k. \tag{2.31}$$

2.4.2.3 Criterion of comparison

The criterion used to compare the performance of the imager was the Peak-Signal-to-Noise-Ratio (PSNR), a common metric implemented in hyperspectral imaging, defined as

$$20 \log_{10} \left(\frac{\max_I}{\mathrm{MSE}^{1/2}} \right), \tag{2.32}$$

where $\max_I$ is the maximum possible value of the image and MSE is the mean squared error with respect to the ground-truth.

2.4.2.4 Sparsifying basis and iterative algorithm

The hyperspectral image was recovered by solving the optimization problem in Eq. (1.2) using the *Gradient Projection for Sparse Reconstruction Algorithm* (GPSR) [33]. The sparsity basis was chosen as $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{DCT} \otimes \boldsymbol{\Psi}_{W}$, where $\boldsymbol{\Psi}_{W}$ is the 2D Wavelet

Symlet 8 basis, $\mathbf{\Psi}_{DCT}$ is the Discrete Cosine basis and $\otimes$ is the Kronecker product [34]; this representation basis has been already used for other imagers such as CASSI [14]. The value of τ in Eq. (1.2) was chosen such that an optimal PSNR is reached.

2.4.2.5 Simulation results

Figure 2.15-left shows the PSNR of the reconstructed scene for the different imagers as a function of the number of captured shots. The SSCSI outperforms conventional CASSI and Colored-CASSI with 4 filters, while exhibiting similar performance to the ideal Colored-CASSI. The RGB profiles of the ground-truth and the reconstructed hyperspectral scenes for $Q = 3$ snapshots can be seen in Fig. 2.15-right. The final spatial quality given by the SSCSI can be inferred from the figures; this can be seen as well in the zoomed portion of the scenes, shown in Fig. 2.16-left. The reconstructed spectral signatures for two different pixels T_1 and T_2 can also be seen in Fig. 2.16-right. As depicted, the SSCSI, the Ideal Colored-CASSI, and the Colored-CASSI with 4 filter throw spectrally accurate results.

2.4.3 Spatially and spectrally super-resolved scenes

2.4.3.1 Spatial super-resolution

As previously mentioned, if the condition $\Delta_c/(1-s) \leq \Delta_d$ holds, it is possible to recover a spatially super-resolved hyperspectral scene, with dimensions specified by Eq. (2.20), and following the model given by Eq. (2.19). This scenario was simulated by defining $\Delta_c = \Delta_d/2$, and a detector size of $N_d \times N_d$ with $N_d = 128$; to recover 24 spectral bands, and assuming that $\beta = 1$, the coded aperture must be located at $s \approx 0.1$. With all the defined parameters, a spectral scene with dimensions of $233 \times 256 \times 24$ can be recovered from the captured measurements. Reconstruction results for a compression ratio of $CR = 0.68$, where the compression ratio is defined as $CR = \frac{QN_d^2}{N_x N_y L}$, can be seen in Fig. 2.17-top, while the equivalent low-resolution scene of $128 \times 128 \times 24$ can be seen in Fig. 2.17-bottom; notice how, in the zoomed portion, the spatial details are clearly defined in the super-resolved case. The PSNR of the recovered super-resolved

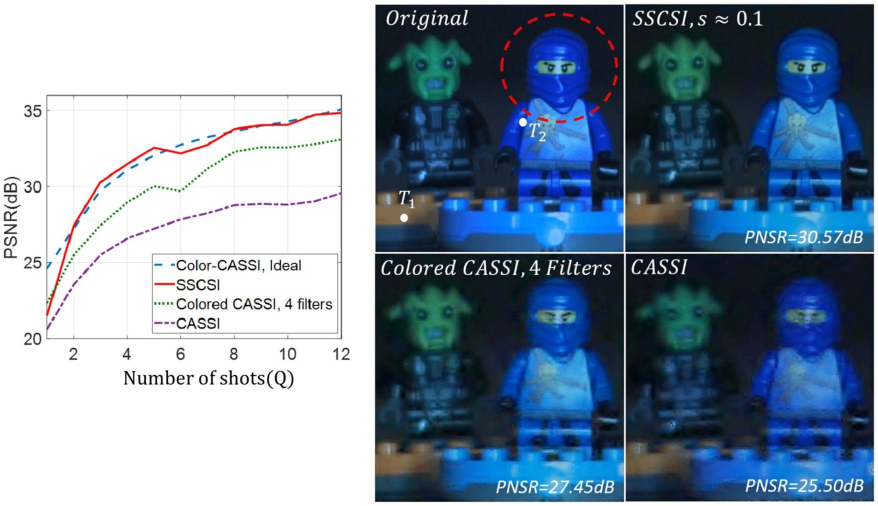

Figure 2.15: Left: PSNR of the reconstruction as a function of the number of shots for the SSCSI, $s \approx 0.1$, CASSI, Colored-CASSI with 4 filters and Ideal Colored-CASSI. All the reconstructions were done assuming $\Delta_c = \Delta_d$. Right: RGB profiles of the original and recovered datacubes. Top-left: Original scene. Top-right: SSCSI reconstruction, $s \approx 0.1$. Bottom-left: Colored-CASSI with 4 filters reconstruction. Bottom-right: CASSI reconstruction. The simulations were done assuming $Q = 3$ snapshots and $\Delta_c = \Delta_d$. The dashed red circle indicates the area to be zoomed in Fig. 2.16.

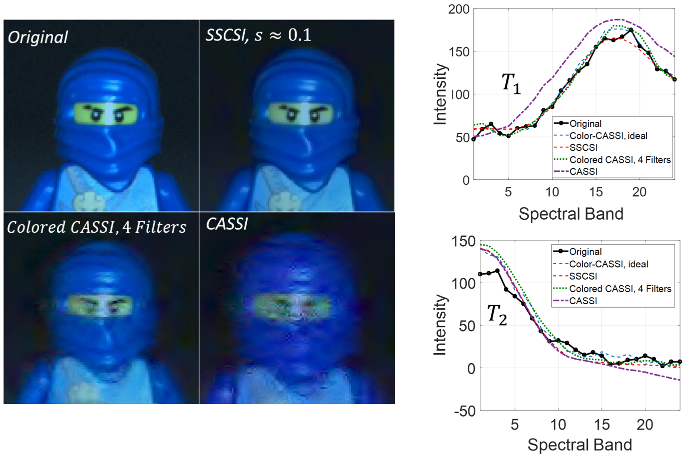

Figure 2.16: Left: zoomed portion of the RGB profiles indicated by the dashed red circle in Fig. 2.15. Right: Spectral signatures for pixels T_1 and T_2. T_1 and T_2 are specified in Fig. 2.15.

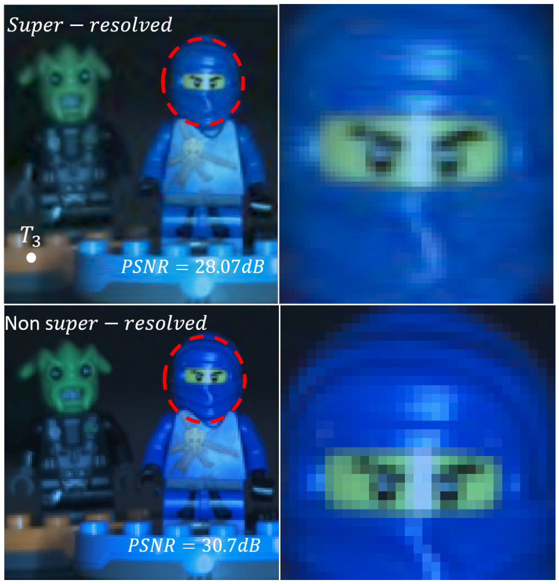

Figure 2.17: Top left: RGB profile of the super-resolved $233 \times 256 \times 24$ datacube, with $s \approx 0.1$, $\Delta_c = \Delta_d/2$, and $CR = 0.68$. Top right: Zoomed portion of the datacube. Bottom left: RGB profile of the non-super-resolved datacube, with $s \approx 0.1$ and $\Delta_c = \Delta_d$. Bottom right: Zoomed portion of the datacube indicated by the dashed red circles.

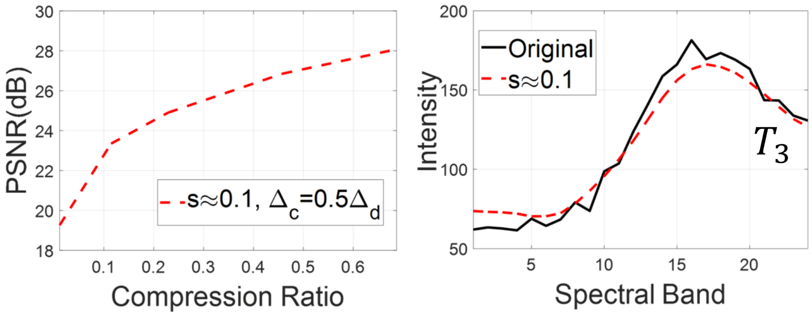

Figure 2.18: Left: PSNR of the super-resolved datacube as a function of the compression ratio, with $s \approx 0.1$ and $\Delta_c = \Delta_d/2$. Right: Recovered spectral signature of T_3, for the super-resolved datacube and $CR = 0.68$. T_3 is specified in Fig. 2.17.

scene for different compression ratios can be seen in Fig .2.18-left, while the recovered spectral signature for pixel T_3 can be seen in Fig. 2.18-right.

2.4.3.2 Spectral super-resolution

Unlike CASSI, the resolution limits on SSCSI can be modified by simply moving the coded aperture on the imager. The number of spectral bands is directly proportional to s, as seen in Eq. (2.18), reaching a maximum when $s = 1$. In other words, a spectral zooming operation can be done by displacing the mask towards the spectral plane. To test that, a spectral signature with strong peaks, synthetically generated, was randomly located in the original datacube, and the reconstructions were performed, assuming that $\beta = 1$, $\Delta_c = \Delta_d$, and the compression ratio was defined as $CR = 0.2$. Notice in Fig. 2.19 that as s increases, the peaks on the spectral signature are clearer; this is quantified by the absolute value of the correlation $|r|$, between original and reconstructed synthetic signatures, which can also be seen in Fig. 2.19.

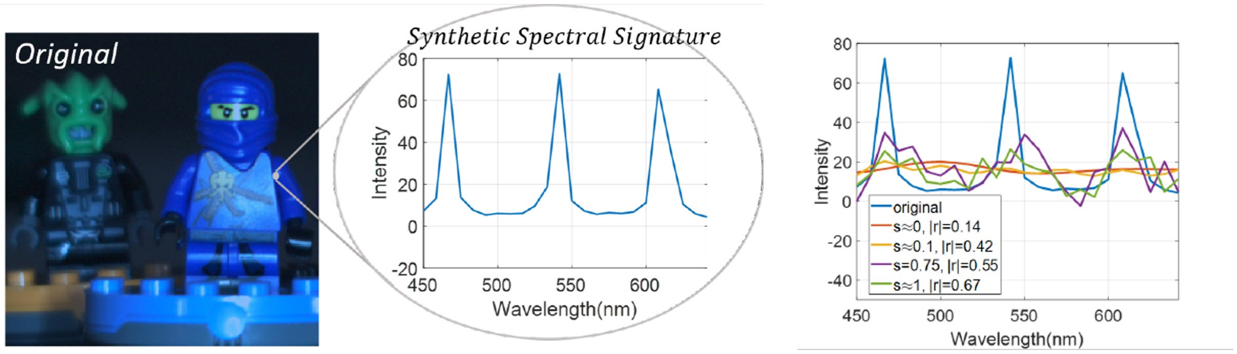

Figure 2.19: Left: Synthetic spectral signature created to evaluate the influence of the parameter s on the spectral resolution. Right: Recovered synthetic spectral signature for different values of the parameter s and a compression ratio of $CR = 0.2$. The variable $|r|$ indicates the absolute value of the correlation coefficient between the original and the reconstructed signatures.

2.4.4 SSCSI performance at different values of s

The SSCSI performance is extremely dependent on s; both the spatial and spectral resolution are defined by this parameter; thus, the final quality of the recovered scene is also affected by s. In order to test that, recovery of the same hyperspectral scene of size $256 \times 256 \times 24$ was attempted with a value of s smaller than that indicated by Eq. (2.18), ($s \approx 0.1$), when $\Delta_c = \Delta_d$ and $\beta = 1$. Results can be seen in Fig. 2.20; as depicted, when $s \approx 0$, colors on the scene are not distinguishable. The PSNR of the reconstructed scene as a function of s can be seen in Fig, 2.21-left, while the reconstructed signature for pixel T_4 can be seen in Fig. 2.21-right; notice that as s approaches to 0.1, the spectral details are better reconstructed (see zoomed portion of the spectral signature).

Figure 2.20: Original and reconstructed RGB profiles for 3 different coded aperture positions, $Q = 6$ snapshots and $\Delta_c = \Delta_d$. The codes exhibit Boolean structure. The optimal PSNR occurs at $s = 0.02$, while at $s \approx 0$ the SSCSI does not recover spectral information.

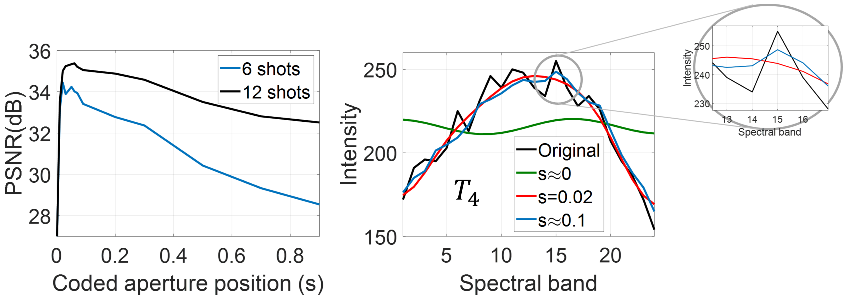

Figure 2.21: Left: PSNR of the reconstructed $256 \times 256 \times 24$ hyperspectral scene as a function of the coded aperture position s, for $Q = 6$ and $Q = 12$ snapshots. The codes exhibit Boolean structure. Right: Recovered signature for pixel T_4 and different values of s. The zoomed portion emphasizes the difference between the signatures and how the spectral details are lost for $s < 0.1$. T_4 is specified in Fig. 2.20.

2.5 Experimental reconstructions

The SSCSI was mounted in an optical table following the indications given by Lin et. al. in the supplementary material for [17]. The experimental setup can be seen in Fig. 2.22. The objective lens, TAMRON-AF 70-300mm, images the scene onto the 300 grooves/mm transmissive diffraction grating; after that, the optical path bends according to the diffraction angle, and the diffracted light passes through a 4-F system composed of two 75mm, 2" lenses. The 4F system creates an image plane prior to the sensor, allowing access to coded aperture positions close to $s = 0$. Then, the light is coded using a lithography-fabricated coded aperture, and that coded information is imaged into a CCD monochrome scientific camera using a 35mm, 1" lens.

2.5.1 Spectral resolution analysis

The theoretical spectral resolution of the SSCSI can be calculated by determining the value of β in the laboratory; it was found that the spectral plane composed

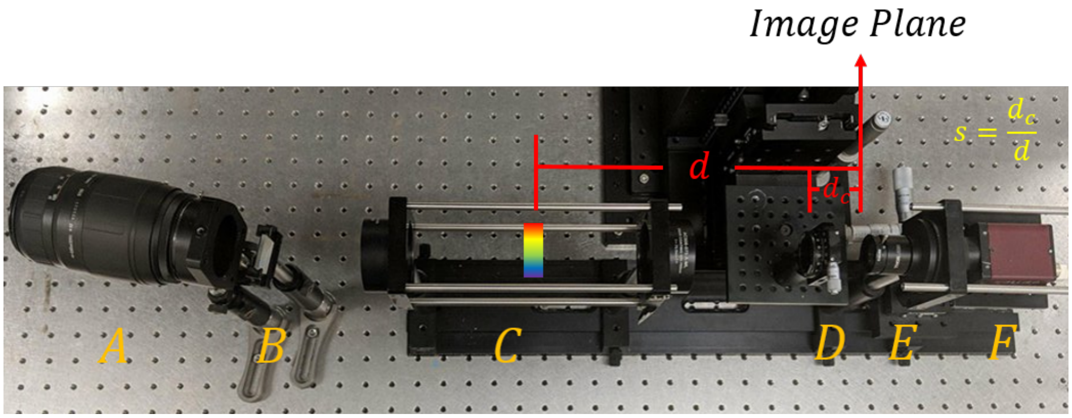

Figure 2.22: SSCSI experimental set up. 6 main components are distinguished. A. TAMRON AF 70-300mm objective lens. B. 300 grooves/mm transmissive diffraction grating. C. 4f system composed of two 75mm, 2” lenses. D. Coded aperture. E. Relay system composed of a 35mm, 1” lens. F. Stingray$^{\text{TM}}$ 640×480 CCD monochrome camera with $9.9\mu m$ pitch size. As depicted, the 4f system allows more flexibility when displacing the coded aperture. Notice that the optical arm is bent according to the diffraction grating angle.

of frequencies from $480nm$ to $620nm$ has a length three times bigger than the implemented coded aperture width; therefore, $\beta \approx 3$; this was determined by locating a white board on the spectral plane and measuring the physical distance between $480nm$ and $620nm$. The number of resolvable bands, assuming $\Delta_c = \Delta_d$ and $N_d = 256$, is equal, according to Eq. (2.18), to $L = \lceil 256s\beta \rceil$. The experimental spectral resolution, on the other hand, was determined using the criterion depicted in Fig. 2.23; two adjacent spectral bands will be resolvable if the projected coded aperture onto the sensor for the bands are shifted by one column. This can be tested by comparing a randomly selected row of the patterns. Table 2.4 shows the experimental and theoretical spectral resolution for different values of s.

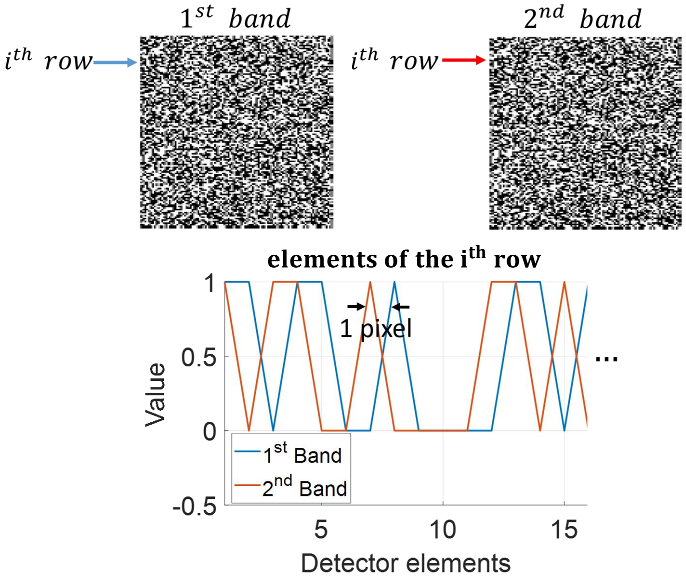

Figure 2.23: Experimental calculation of the spectral resolution. Top: effective coded aperture for two consecutive bands. Bottom: Comparison of i^{th} effective coded aperture row elements for the two consecutive bands.

Table 2.4: Theoretical and experimental spectral resolution for different s parameters.

s	Δ_λ, Theoretical (nm)	Δ_λ, Experimental (nm)
0.004	43	40
0.0078	22	28
0.011	16	20

2.5.2 Experimental datacube reconstruction

A hyperspectral scene was reconstructed using the SSCSI and two snapshots. The implemented coding patterns exhibit a Boolean structure and have a pixel pitch size of $\Delta_c = 19.4\mu m$. In the calibration process, the sensing matrix $\mathbf{H}$ was assembled by locating a white board as a scene and then capturing the coding pattern per each spectral band to be recovered; this has to be done for the two captured snapshots.

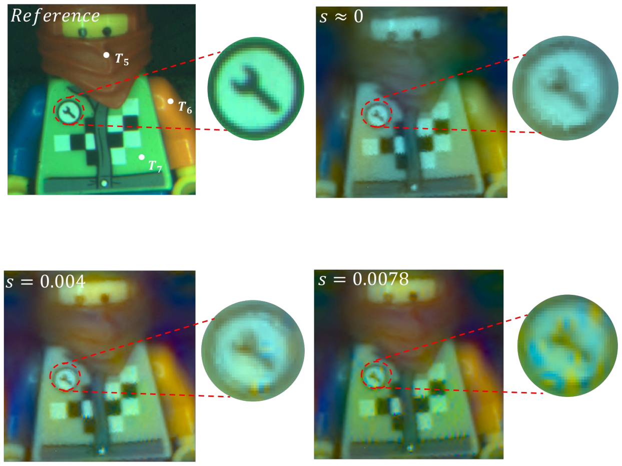

Figure 2.24: RGB profiles of the reference scene and the reconstructed datacubes for three different values of s and two complementary shots. $s \approx 0$, $s = 0.004$, and $s = 0.0078$.

After that, each pattern is located in its respective diagonal of $\mathbf{H}$. Once the calibration is done, the white board is replaced by the target. Then, the scene is properly illuminated and the coded scene is captured for each shot, to then assemble the measurements vector. Then, the reconstruction algorithm is run. The RGB-mapping of the reconstructed spectral scene for different values of s can be seen in Fig. 2.24. As depicted, as s increases, the spectral information becomes more noticeable; this can also be seen in Figs. 2.25 and 2.26. Nevertheless, as s increases, some artifacts start to appear in the reconstruction; this will be discussed in the next subsection.

2.5.3 SSCSI experimental limitations

All the presented discrete measurements models rely on the assumption that the aperture of the objective lens is infinitesimally small; this is an ideal case that cannot be replicated in experiments for several reasons. First, the light throughput of the system is extremely affected by the aperture of the objective lens. In addition, the minimum resolvable spatial feature of any optical system is determined by the diffraction limit, which is proportional to the lens aperture. In experiments, the aperture must have a finite length; this finite aperture introduces blurring and artifacts into the final

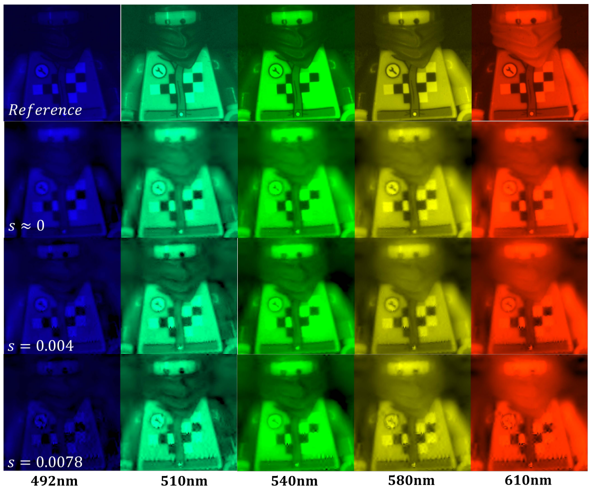

Figure 2.25: Reference and reconstructed spectral bands for three different values of s and two complementary shots. From top to bottom: Reference, $s \approx 0$, $s = 0.004$, and $s = 0.0078$.

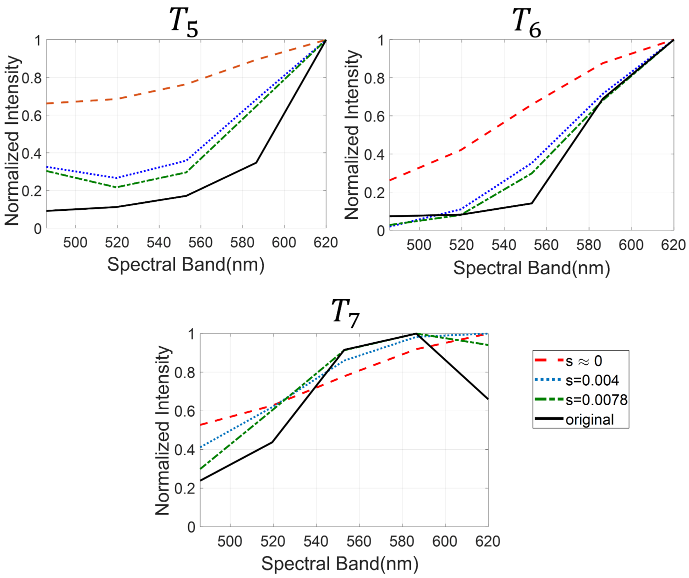

Figure 2.26: Original and reconstructed spectral signatures for pixels T_5, T_6, and T_7 and three different values of s, $s \approx 0$, $s = 0.004$, and $s = 0.0078$. T_5, T_6, and T_7 are specified in Fig. 2.24.

reconstruction, which are directly proportional to the value of s. Other limitation in SSCSI comes from the spectral plane length; as found in the experiments, the spectral plane length is three times bigger than the coded aperture width, which leads the extreme wavelengths (blue and red), to impinge out of the mask and be discarded. That is why the range of interest was adjusted in the experiment to $[480nm - 620nm]$.

2.6 Conclusions

In this chapter, a rigorous analysis of the SSCSI spatial and spectral resolution limits is proposed and developed. It was found that the minimum attainable feature in space and spectrum depend on parameters such as the dispersion introduced by the grating, the coded aperture and detector pixel pitch sizes, and the coded aperture position with respect to the sensor. Based on these parameters, two discrete measurements models were defined. In the first one, when $\Delta_c/(1-s) \leq \Delta_d$, a spatially super-resolved scene can be recovered from a low resolution sensor. The minimum recoverable spatial feature is defined by the region $\Delta_c/(1-s) \times \Delta_c$, and the discrete measurements model is defined by Eq. (2.19), where the auxiliary variable $W_{m,m'}$ accounts for the mismatching between detector and code aperture. In the second one, when $\Delta_c/(1-s) > \Delta_d$, the minimum recoverable spectral feature is defined by $\Delta_d \times \Delta_c$; therefore, it is possible to recover a spatially super-resolved scene in one dimension. The discrete measurements model is given by Eq.(2.23), where the variable p is used to create a synthetic gray-scale coded aperture.

An upper-bound of the SSCSI spectral resolution was defined in theory and experiments as $\Delta_c/s\alpha$; hence, a zooming operation over the spectral domain can be done in SSCSI by moving the coded aperture towards the spectral plane, reaching a maximum of resolvable bands when $s = 1$ (spectral plane).

The SSCSI was mounted and tested in experiments to test the spectral resolution dependency on s. In this process, several issues, which were not considered in the theoretical model, were found, such as the dependency of the SSCSI performance on the size of the objective lens aperture, the reduced spectral range coming from the extreme wavelengths impinging out of the coded aperture, and the low light throughput of the system.

This is the first research work that fully characterizes the SSCSI in theory and experiments, and, as a result, two conference papers with preliminary results in *Computational Optics, Sensing and Imaging* (Salazar et. al. [35, 36]), and one journal paper in *IEEE transactions on Computational Imaging* (Salazar et. al. [37]) were published (see attached disclaimer for copyright issues). In the next chapter, and based on the proposed model, an optimization framework to find the best coding pattern is proposed and developed.

All the data and software for this chapter are available in [38].

Chapter 3

CODED APERTURE OPTIMIZATION IN THE SPATIAL SPECTRAL COMPRESSIVE SPECTRAL IMAGER

3.1 Introduction

The coding process in Compressive Spectral Imagers is an essential step directly influencing the final quality of the recovered information. Although fully random patterns offer good results[39], several efforts have been made to optimize the coded aperture structure for CASSI and Colored-CASSI, taking into account the physics and discretization of the sensing process. One of the first criteria used to determine the sensing matrix quality was the *Restricted Isometry Property* (RIP) [40], which is defined as follows

$$(1-\delta_S)||\boldsymbol{\pi}||_2^2 \leq ||\mathbf{A}\boldsymbol{\pi}||_2^2 \leq (1+\delta_S)||\boldsymbol{\pi}||_2^2, \tag{3.1}$$

where $\mathbf{A}$ is the sensing matrix and δ_S is the Restricted Isometry Constant (RIC). Minimizing δ_S leads to better reconstructions. A closed definition of δ_S is given by

$$\delta_S = \max_{\rho\subset[N^2L],|\rho|\leq S} ||\mathbf{A}_\rho^T\mathbf{A}_\rho - \mathbf{I}||_2, \tag{3.2}$$

where $[N^2L] := [1,2,\ldots N^2L]$, $|\rho|$ is the cardinality of ρ, $\mathbf{I}$ is the identity matrix and $\mathbf{A}_\rho$ is a $N^2Q \times |\rho|$ matrix, with $|\rho|$ columns of $\mathbf{A}$, indexed by ρ. The RIP has been extensively used to find optimal patterns in CASSI; Wagadarikar [41] defined the necessary conditions on the coding patterns for the CASSI sensing matrix to satisfy the RIP; Correa et al. in [42] used the RIP property to find the optimal patterns for CASSI, and their interdependency among shots. Other RIP-based coded aperture optimization algorithms were proposed in [43, 44, 45]. Recently, with the increasing of computational resources, several deep and machine learning-based attempts at finding the optimal patterns for CASSI have been made [46, 47]. In [46], for example, Wang et. al. proposed the *Hyperreconnect* architecture, where through a Convolutional Neural Network, coding process and reconstruction are jointly optimized. In this chapter, an optimization framework for coded aperture patterns in SSCSI is proposed and developed; this method is based on an upper-bound of the RIP and, unlike deep learning techniques, the results are scene independent and merely based on the physics of the sensing process. This chapter is divided as follows: in Section II, the discrete measurements model is rewritten for the special case when $\Delta_d = \Delta_c$; in Section III, the cost function implemented in the optimization is developed in detail; in Section IV, the algorithm to find the optimal coded apertures is explained; in Section V, a new spectral statistics called *Axially Averaged Power Spectral Density*, implemented to characterize the spectral content of the patterns, is proposed and explained; in Section VI, the performance of the optimal coded apertures is tested against conventional patterns; finally, this chapter discusses the conclusions and highlights of this research line.

3.2 Rewriting the discrete measurements model

The proposed optimization algorithm in this chapter is developed for the particular case when the coded aperture and detector pitch sizes are the same, that is $\Delta_c = \Delta_d = \Delta$; with this assumption, m' in Eq. (2.23) can be written as $m' = \left\lfloor m(1-s) + \frac{s\alpha\lambda_{min}}{\Delta} \right\rfloor + 1$, while the variable p_m can be rewritten as

$$p_m = \begin{cases} 1 & \text{If} \quad (m+1)\Delta \leq \frac{m'\Delta - s\alpha\lambda_{min}}{(1-s)} \\ \frac{m'\Delta - s\alpha\lambda_{min}}{\Delta(1-s)} - m & \text{If} \quad (m+1)\Delta > \frac{m'\Delta - s\alpha\lambda_{min}}{(1-s)}. \end{cases} \tag{3.3}$$

If $\alpha\lambda_{min}$ is assumed to be zero, that is, if λ_{min} impinges on the leftmost side of the coded aperture, $m' = \lfloor m(1-s) \rfloor + 1$, and Eq. (2.23) can be rewritten as follows

$$\mathbf{g}_{m,n} = \sum_{k=0}^{L-1} \left(\mathbf{t}_{\lfloor m(1-s) \rfloor + k,n} \times p_m + \mathbf{t}_{\lfloor m(1-s) \rfloor + 1 + k,n} \times (1 - p_m) \right) \mathbf{f}_{m,n,k}, \tag{3.4}$$

where p_m, given in (3.3), is redefined as

$$p_m = \begin{cases} 1 & \text{If} \quad m+1 \leq \frac{\lfloor m(1-s)\rfloor+1}{(1-s)} \\ \frac{\lfloor m(1-s)\rfloor+1}{1-s} - m & \text{If} \quad m+1 > \frac{\lfloor m(1-s)\rfloor+1}{(1-s)}. \end{cases} \tag{3.5}$$

The number of resolvable spectral bands given by Eq. (2.18) can be rewritten as

$$L = \lceil s\beta N \rceil . \tag{3.6}$$

3.3 Coded aperture optimization

The SSCSI sensing matrix is defined as $\mathbf{A} = \mathbf{H}\mathbf{\Psi}$, where $\mathbf{H}$ is the forward model matrix fully described in the last chapter, and $\mathbf{\Psi}$ is the basis implemented to sparsify the data. This matrix determines how accurate the reconstructed scenes will be in terms of spatial and spectral quality; therefore, careful design of $\mathbf{A}$ must be developed as elaborated next.

3.3.1 The minimization problem

Finding the optimal coded aperture patterns using the RIP is equivalent to minimizing δ_s given by Eq. (3.2); however, this problem can be NP-hard [48]; hence an upper-bound of δ_s must be established as follows:

$$\delta_s \leq ||\mathbf{A}^T\mathbf{A} - \mathbf{I}||_F, \tag{3.7}$$

where $||\cdot||_F$ is the Frobenius norm of a matrix. This approximation has already been used for code design in Computed Tomography [49, 50]. Finding the SSCSI optimal patterns through the minimization of δ_s, is a problem that can be written as follows:

$$\begin{gathered} \min_{\mathbf{H}} ||\mathbf{A}^T\mathbf{A} - \mathbf{I}||_F^2 \\ \text{subject to } \mathbf{H} \in \mathbf{C}_{N,L,Q}, \end{gathered} \tag{3.8}$$

where $\mathbf{C}_{N,L,Q}$ represents the set of matrices for the SSCSI system, reconstructing a $N \times N \times L$ datacube, from Q measured snapshots with a $N \times N$ detector. This last optimization problem can be reformulated as follows

$$\begin{aligned} \min_{\mathbf{H}} \quad & \sum_{b,c=1,b\neq c}^{N^2L} |\left(\mathbf{A}^T\mathbf{A}\right)_{b,c}|^2 \\ \text{subject to:} \quad & \mathbf{H} \in \mathbf{C}_{N,L,Q} \\ & \left(\mathbf{A}^T\mathbf{A}\right)_{b,b} = 1 \ \ \forall \ \ b \in [1, \ldots N^2L], \end{aligned} \tag{3.9}$$

where $\left(\mathbf{A}^T\mathbf{A}\right)_{b,c} = \langle \mathbf{A}(:,b), \mathbf{A}(:,c)\rangle$ is the $(b,c)^{th}$ element of $\mathbf{A}^T\mathbf{A}$, and $\langle\cdot\rangle$ represents the inner product operation. Finding the optimal matrix $\mathbf{H}$, and the corresponding coded aperture pattern requires first an extensive analysis of the structure of $\mathbf{H}$, as shown next.

3.3.2 $\mathbf{A}^T\mathbf{A}$ as a function of t

Given that $\mathbf{A} = \mathbf{H}\mathbf{\Psi}$, $\left(\mathbf{A}^T\mathbf{A}\right)_{b,c} = \langle \mathbf{A}(:,b), \mathbf{A}(:,c)\rangle$ can be rewritten as follows [51]:

$$\left(\mathbf{A}^T\mathbf{A}\right)_{b,c} = \sum_{i,j=1}^{N^2L} \langle \mathbf{h}_i, \mathbf{h}_j\rangle \, \boldsymbol{\psi}_i(b)\boldsymbol{\psi}_j(c), \tag{3.10}$$

where $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots \mathbf{h}_{N^2L}]$ and $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots \boldsymbol{\psi}_{N^2L}]$. The term $\langle \mathbf{h}_i, \mathbf{h}_j\rangle$, is different from zero, if and only if i and j are separated by an integer multiple of N^2, or $j = \left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i$, where $k = [0, 1, \ldots L-1]$. This comes from analyzing the structure of $\mathbf{H}$, as depicted in Fig. 3.1. Eq. (3.10) can be rewritten as follows

$$\left(\mathbf{A}^T\mathbf{A}\right)_{b,c} = \sum_{i=1}^{N^2L}\sum_{k=0}^{L-1} \left\langle \mathbf{h}_i, \mathbf{h}_{\left(k-\left\lfloor \frac{i-1}{N^2}\right\rfloor\right)N^2+i} \right\rangle \times \boldsymbol{\psi}_i(b)\boldsymbol{\psi}_{\left(k-\left\lfloor \frac{i-1}{N^2}\right\rfloor\right)N^2+i}(c). \tag{3.11}$$

Likewise, the vector $\mathbf{h}_i$ can be expressed as $\mathbf{h}_i = [\hat{\mathbf{t}}^{m_i,n_i,\ell_i,1}, \hat{\mathbf{t}}^{m_i,n_i,\ell_i,2}, \ldots, \hat{\mathbf{t}}^{m_i,n_i,\ell_i,Q}]^T$, where $\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q} = \mathbf{t}_{\lfloor m_i(1-s)\rfloor+\ell_i,n_i,q} \times p_{m_i} + \mathbf{t}_{\lfloor m_i(1-s)\rfloor+1+\ell_i,n_i,q} \times (1-p_{m_i})$, p_{m_i} is defined in Eq. (3.5), and $\mathbf{t}_{m,n,q}$ is the pixel of the coded aperture pattern located at (m,n) and shot q. The variables m_i, n_i and ℓ_i can be defined in terms of i, as $m_i = \left\lfloor \frac{i-1}{N} \right\rfloor - \left\lfloor \frac{i-1}{N^2} \right\rfloor N$,

$n_i = \mathrm{mod}(i-1, N)$ and $\ell_i = \left\lfloor \frac{i-1}{N^2} \right\rfloor$. It can be proven that, when $j = \left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i$, $m_j = m_i$, $n_j = n_i$ and $\ell_j = k$ (see appendix C). Taking into account the above relationships, the term $\left\langle \mathbf{h}_i, \mathbf{h}_{\left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i} \right\rangle$ can be written as

$$\left\langle \mathbf{h}_i, \mathbf{h}_{\left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i} \right\rangle = \sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i, n_i, \ell_i, q} \hat{\mathbf{t}}^{m_i, n_i, k, q}. \tag{3.12}$$

Replacing Eq. (3.12) into Eq. (3.11) leads to

$$\left(\mathbf{A}^T \mathbf{A}\right)_{b,c} = \sum_{i=1}^{N^2 L} \sum_{k=0}^{L-1} \sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i, n_i, \ell_i, q} \hat{\mathbf{t}}^{m_i, n_i, k, q} \times \boldsymbol{\psi}_i(b) \boldsymbol{\psi}_{(k-\ell_i) N^2 + i}(c). \tag{3.13}$$

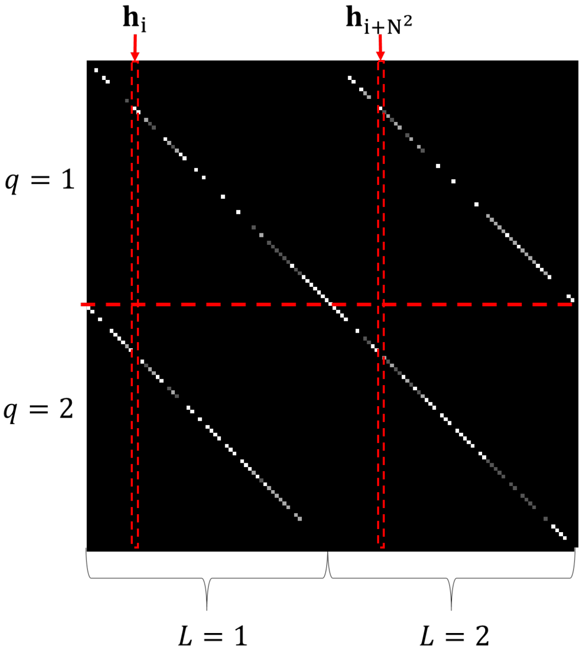

Figure 3.1: Structure of the matrix $\mathbf{H}$ for $Q = 2$ snapshots, $L = 2$ spectral bands and $s = 0.25$. Notice that $\langle \mathbf{h}_i, \mathbf{h}_j \rangle \neq 0$ if i and j are separated by a integer multiple of N^2, or $j = \left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i$.

3.3.3 Minimizing the off-diagonal elements

In order to minimize the term $\sum_{b,c=1, b\neq c}^{N^2 L} |\left(\mathbf{A}^T \mathbf{A}\right)_{b,c}|^2$ in Eq. (3.9), one can consider the term $\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i, n_i, \ell_i, q} \hat{\mathbf{t}}^{m_i, n_i, k, q} \times \boldsymbol{\psi}_i(b) \boldsymbol{\psi}_{(k-\ell_i) N^2 + i}(c)$, from Eq. (3.13) as the one to be minimized. Given that $\boldsymbol{\psi}_i(b)\boldsymbol{\psi}_{(k-\ell_i)N^2+i}(c)$ are fixed quantities dictated by the sparsifying basis, the problem is reduced to analyzing the term $\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\hat{\mathbf{t}}^{m_i,n_i,k,q}$. Notice that the variable k can take values from 0 to $L-1$; this means that, when searching for the minimum of $\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\hat{\mathbf{t}}^{m_i,n_i,k,q}$, one needs to consider the coded aperture elements on a neighborhood of length L, over a given row of the coded aperture n_i, independently for each snapshot. Intuitively, this minimization problem can be seen as separating the coded aperture pixels that are set to one, as much as possible, over a row of the pattern. A similar constraint also appears when the sensing matrix of CASSI is analyzed [42, 44].

3.3.4 On-diagonal elements of $\mathbf{A}^T\mathbf{A}$

For $b = c$, Eq. (3.13) can be rewritten as the summation of two terms

$$\begin{aligned}\left(\mathbf{A}^T\mathbf{A}\right)_{b,b} &= \sum_{i=1}^{N^2L}\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 \times \boldsymbol{\psi}_i(b)^2 + \\ &\quad \sum_{i=1}^{N^2L}\sum_{\substack{k=0\\k\neq\ell_i}}^{L-1}\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\hat{\mathbf{t}}^{m_i,n_i,k,q} \times \boldsymbol{\psi}_i(b)\boldsymbol{\psi}_{(k-\ell_i)N^2+i}(b). \qquad (3.14)\end{aligned}$$

Notice that the second term in Eq. (3.14) will tend to zero as a consequence of minimizing $\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\hat{\mathbf{t}}^{m_i,n_i,k,q}$, which was mentioned in the last subsection. Hence, to reach the condition $\left(\mathbf{A}^T\mathbf{A}\right)_{b,b} = 1$, one must find a set of coded aperture patterns such that $\sum_{i=1}^{N^2L}\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 \times \boldsymbol{\psi}_i(b)^2 \approx 1$. Let $\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2$ be a constant number C for any value of m_i, n_i, and ℓ_i; this leads to rewriting the expression $\sum_{i=1}^{N^2L}\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 \times \boldsymbol{\psi}_i(b)^2$, as $C\sum_{i=1}^{N^2L} \boldsymbol{\psi}_i(b)^2$. Given that the basis $\boldsymbol{\Psi}$ is usually chosen to be an orthonormal matrix, $\sum_{i=1}^{N^2L} \boldsymbol{\psi}_i(b)^2 = 1$, and $C\sum_{i=1}^{N^2L} \boldsymbol{\psi}_i(b)^2 = C$. The analysis above affirma that, if one guarantees that $\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 = 1$, or $||\mathbf{h}_i||_2^2 = 1$, the term $\left(\mathbf{A}^T\mathbf{A}\right)_{b,b}$ would be equal to 1. The condition $\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 = 1$ holds, if $\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q} = 1$ for $q = q_o$ and $\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q} = 0$ for $q \neq q_o$. In order to set $\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q} = \mathbf{t}_{\lfloor m_i(1-s)\rfloor+\ell_i,n_i,q} \times p_{m_i} + \mathbf{t}_{\lfloor m_i(1-s)\rfloor+1+\ell_i,n_i,q} \times (1-p_{m_i})$ to 0, the elements of the coded aperture $\mathbf{t}_{\lfloor m_i(1-s)\rfloor+\ell_i,n_i,q}$ and $\mathbf{t}_{\lfloor m_i(1-s)\rfloor+1+\ell_i,n_i,q}$ must both be set to 0. Likewise, to set

$\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q_o}$ to 1, $\mathbf{t}_{\lfloor m_i(1-s)\rfloor+\ell_i,n_i,q_o}$ and $\mathbf{t}_{\lfloor m_i(1-s)\rfloor+1+\ell_i,n_i,q_o}$ must both be set to 1. This means that two different constraints must be considered to make $\left(\mathbf{A}^T\mathbf{A}\right)_{b,b}$ equal to 1; the first is the complementarity condition among the captured snapshots, which can be expressed as $\sum_{q=1}^{Q}\mathbf{t}_{m,n,q} = 1 \;\forall\; m,n$. The second is a clustering condition that makes two consecutive elements of the coded aperture, $\mathbf{t}_{\lfloor m_i(1-s)\rfloor+\ell_i,n_i,q_o}$ and $\mathbf{t}_{\lfloor m_i(1-s)\rfloor+1+\ell_i,n_i,q_o}$, over the row n_i, to be 1.

3.3.5 The cost minimization problem

In order to find the optimal codes to satisfy the previously mentioned conditions, the following cost function is proposed

$$\mathbf{t}_{opt} = \min_{\mathbf{t}} \left(\sum_{i=1}^{N^2L} \sum_{\substack{k=0\\k\neq\ell_i}}^{L-1} \left\langle \mathbf{h}_i, \mathbf{h}_{(k-\ell_i)N^2+i} \right\rangle^2 + c_1 \sum_{i=1}^{N^2L} \left| ||\mathbf{h}_i||_2^2 - 1 \right|^2 \right)$$

$$\text{subject to:} \quad \textstyle\sum_{q=1}^{Q}\mathbf{t}_{m,n,q} = 1 \;\forall\; m,n. \tag{3.15}$$

By using Eq. (3.12), the last expression can be rewritten as follows

$$\mathbf{t}_{opt} = \min_{\mathbf{t}} \left(\sum_{i=1}^{N^2L} \sum_{\substack{k=0\\k\neq\ell_i}}^{L-1} \left(\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\hat{\mathbf{t}}^{m_i,n_i,k,q} \right)^2 + c_1 \sum_{i=1}^{N^2L} \left| \sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 - 1 \right|^2 \right)$$

$$\text{subject to:} \quad \textstyle\sum_{q=1}^{Q}\mathbf{t}_{m,n,q} = 1 \;\forall\; m,n. \tag{3.16}$$

The first term is related to the minimization of the off-diagonal elements, while the second term is related to the constraint $\left(\mathbf{A}^T\mathbf{A}\right)_{b,b} = 1 \;\forall\; b \in [1,\ldots,N^2L]$. The variable c_1 defines the influence of this constraint on the minimization problem, and must be experimentally tuned to get optimal reconstructions, taking values between 1 to 70. The complementarity condition is imposed by the constraint $\sum_{q=1}^{Q}\mathbf{t}_{m,n,q} = 1 \;\forall\; m,n$. Notice that Eq. (3.16) represents a relaxed version of the optimization problem given by Eq. (3.9) that avoids the explicit calculation of the sparsifying basis, since this requires a high computational effort [50].

3.3.6 Separating the ON pixels through the y axis

The separation of the coded aperture pixels that are set to one over the different columns has been proven to strongly influence the quality of the final recovered scene on different compressive spectral imagers, such as CASSI [42, 44]. Correa et al. [42] proved that separating the ON pixels over the columns may decrease the coherence of the sensing matrix, and therefore, enhance the reconstructions. Besides, since the SSCSI bases its functionality on the location of the coded aperture in an out-of-focus position, separating the ON pixels prevents the spatial-spectral from overlapping. Therefore, a new condition must be added to Eq. (3.16), as follows,

$$\mathbf{t}_{opt} = \min_{\mathbf{t}} \left(\sum_{i=1}^{N^2L} \sum_{\substack{k=0 \\ k\neq\ell_i}}^{L-1} \left(\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i,n_i,\ell_i,q} \hat{\mathbf{t}}^{m_i,n_i,k,q} \right)^2 \right.$$
$$\left. + c_1 \sum_{i=1}^{N^2L} \left| \sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q} \right)^2 - 1 \right|^2 + \sum_{m,n,q} \Omega^{m,n} \odot \tilde{\mathbf{t}}^{q} \right)$$

$$\text{subject to:} \quad \textstyle\sum_{q=1}^{Q} \mathbf{t}_{m,n,q} = 1 \;\forall\; m,n, \tag{3.17}$$

where $\Omega^{m,n}$ is $2L-1 \times 2L-1$ filter that penalizes the closeness between coded aperture elements set to one, in all directions, $\tilde{\mathbf{t}^{q}}$ is a portion of the q^{th} coded aperture with size $2L-1 \times 2L-1$, centered in m, n, and $\odot$ represents the element-wise multiplication between $\Omega^{m,n}$ and $\tilde{\mathbf{t}}$, and the posterior summation of the resultant elements. The entries of Ω are defined as follows

$$\Omega^{m,n}(a_1, a_2) = \begin{cases} 0 & \text{if} \quad a_1, a_2 = m, n \\ \frac{1}{\sqrt{(a_1-m)^2+(a_2-n)^2}} & \text{otherwise.} \end{cases} \tag{3.18}$$

Notice that, as a_1, a_2 are further from m, n, the magnitude of the coefficients of $\Omega^{m,n}$ decreases. The spatial size of $\Omega^{m,n}$ and $\tilde{\mathbf{t}}$ might be less than $2L-1 \times 2L-1$, if m, n are close to 0 or $N-1$. Figure 3.2 shows the structure of Ω and the process to calculate $\Omega \odot \tilde{\mathbf{t}^{q}}$. A similar filtering process, taking into account the frequency response of the human eye, was already proposed by Rueda et al. in [52] for Colored-CASSI using a time-of-flight sensor.

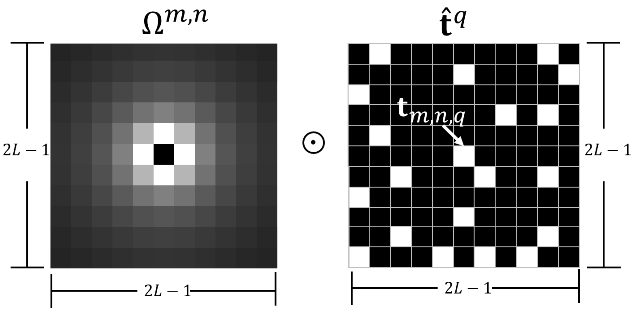

Figure 3.2: Ω filter implemented to penalize the closeness of the ON pixels, on the optimal coded apertures.

3.4 Algorithm implementation

Algorithm 1 shows the steps used to attain optimal coded aperture patterns, guided by Eq. (3.17). The initial coded aperture exhibits a complementary structure (line 1), where again, the term complementary is defined as having just one coded aperture pixel set to 1 at each spatial position of the ensemble of patterns. Besides, the ON pixels in each coded aperture are inversely proportional to the number of captured snapshots. These types of patterns are also referred to as Boolean [32]. The algorithm starts by calculating the value of the cost function given in Eq. (3.17) (line 2). Then, a random walk over the coded aperture elements is done (line 4). For each coded aperture pixel position, the algorithm evaluates the effect of setting to 1 the pixel on a given snapshot, on the cost function (lines 7, 8, 9 and 10). This approach was already explored for colored CASSI [44]. The function **Find**$(r1, r2)$ finds the indexes of i that are affected by $\mathbf{t}_{r1,r2}$. This is done in order to simplify the number of operations in the calculation of the cost function (line 10). After going over all the snapshots, the values of q for which the cost function reaches a minimum, are determined; from those values of q, the algorithm randomly picks one and sets that pixel to 1, while the others are set to 0 (lines 12 and 13). This last constraint forces the complementary condition to hold in any case. The objective function is again calculated (line 15) and the stop criteria ($\epsilon > \epsilon_o$) is evaluated to determine whether the algorithm continues or not. In this paper, ϵ_o is assumed to be 0.1% of the initial value of the cost function.

3.5 Spectral content of coded apertures

The optimal coding structures for CASSI and Colored-CASSI have been shown in theory and experiments to exhibit a pseudo-random structure known as *blue noise* [42, 53, 54, 55]. Ulichney [56] defined a blue noise pattern as one where the white pixels are distributed as homogeneously as possible through the mask; therefore, the spectral content of such pattern is entirely composed of high frequency components. *Green noise*, which will be used here to characterize the SSCSI optimal patterns, was first defined by Lau et. al. [57, 58, 59], as a pattern where clusters of minority pixels are homogeneously distributed over the mask. It is called green noise because its spectral content shows components on middle frequencies. To characterize the different patterns, it is possible to use either spatial or spectral statistics. These metrics describe the distribution of the minority pixels over the pattern. In this dissertation, a spectral statistic called *Axially Averaged Power Spectral Density*, ($AAPSD$), is proposed in order to marginally characterize the patterns over the rows and the columns of the coded aperture.

3.5.1 Axially Averaged Power Spectral Density

Let $\mathbf{t}$ be a coded aperture array, and $\mathcal{DFT}$ be the 2-dimensional Fourier Transform. The power spectral density of that pattern can be defined as follows

$$\mathcal{P}_f = \frac{1}{R}\sum_{i=1}^{R}\frac{|\mathcal{DFT}(\phi_i)|^2}{\mathcal{E}(\phi_i)}, \tag{3.19}$$

where R represents the number of subregions in which the coded aperture is divided, $\mathcal{DFT}(\phi_i)$ is the 2D-Fourier Transform of the subregion ϕ_i, and $\mathcal{E}(\phi_i)$ is the energy of that subregion, which is on the denominator to obtain a normalized spectrum. Notice that the dimensions of ϕ_i will be equal to $\frac{N}{\sqrt{R}} \times \frac{N}{\sqrt{R}}$, and therefore the final dimensions

Algorithm 1: Implemented algorithm to find the optimal coded apertures for SSCSI.

```
Input   : $s$,$N$,$c_1$
Output: $\mathbf{t}_{opt}$
$\mathbf{t} \leftarrow \text{Boolean}(N, Q)$
$F_{old} = \sum_{i=1}^{N^2L} \sum_{\substack{k=0\\k\neq\ell_i}}^{L-1} \left(\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\hat{\mathbf{t}}^{m_i,n_i,k,q}\right)^2 + c_1 \sum_{i=1}^{N^2L} \left|\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 - 1\right|^2 + \sum_{m,n,q} \Omega^{m,n} \odot \tilde{\mathbf{t}}^q$
while $\epsilon > \epsilon_o$ do
    $R \leftarrow \text{2DRandomWalk}(N, N)$
    for $r_1, r_2 \in R$ do
        $\hat{\mathbf{t}} = \mathbf{t}$
        for $q$=1:$Q$ do
            $\hat{\mathbf{t}}_{r_1,r_2,1:Q} = 0$
            $\hat{\mathbf{t}}_{r_1,r_2,q} = 1$
            $\mathcal{I} = \mathbf{Find}(r_1, r_2)$
            $\hat{F}(q) = \sum_{i\in\mathcal{I}} \sum_{\substack{k=0\\k\neq\ell_i}}^{L-1} \left(\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\hat{\mathbf{t}}^{m_i,n_i,k,q}\right)^2 + c_1 \sum_{i\in\mathcal{I}} \left|\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 - 1\right|^2 + \Omega^{r1,r2} \odot \tilde{\mathbf{t}}^q$
        end
        $\mathbf{t}_{r1,r2,1:Q} = 0$
        $\mathbf{t}_{r1,r2,\text{rand}(\min(\hat{F}))} = 1$
    end
    $F_{new} = \sum_{i=1}^{N^2L} \sum_{\substack{k=0\\k\neq\ell_i}}^{L-1} \left(\sum_{q=1}^{Q} \hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\hat{\mathbf{t}}^{m_i,n_i,k,q}\right)^2 + c_1 \sum_{i=1}^{N^2L} \left|\sum_{q=1}^{Q} \left(\hat{\mathbf{t}}^{m_i,n_i,\ell_i,q}\right)^2 - 1\right|^2 + \sum_{m,n,q} \Omega^{m,n} \odot \tilde{\mathbf{t}}^q$
    $\epsilon = |F_{new} - F_{old}|$
    $F_{old} = F_{new}$
end
```

of $\mathcal{P}_f$ will be equal to $\frac{N}{\sqrt{R}} \times \frac{N}{\sqrt{R}}$. This Power Spectral Density estimator is called Bartlett's Method [60], which relies on dividing the original pattern into subregions called periodograms. The *Axially Averaged Power Spectral Density* over the rows is defined as follows

$$AAPSD_x = \frac{\sqrt{R}}{N} \sum_{m=1}^{\frac{N}{\sqrt{R}}} \mathcal{P}_f(m, :). \tag{3.20}$$

The final summation is divided by $\frac{N}{\sqrt{R}}$, in order to take the average over the rows. Likewise, the *Axially Averaged Power Spectral Density* over the columns is defined as follows:

$$AAPSD_y = \frac{\sqrt{R}}{N} \sum_{n=1}^{\frac{N}{\sqrt{R}}} \mathcal{P}_f(:, n). \tag{3.21}$$

Figure 3.3 shows the process of calculating the power spectral density $\mathcal{P}_f$ for a 16×16 coded aperture using $R = 4$ subregions. Figure 3.4 illustrates the calculation of the *Axially Averaged Power Spectral Density* over the rows and the columns for a given 32×32 power spectral density $\mathcal{P}_f$.

3.6 Simulation results

The proposed algorithm was run to find the SSCSI optimal patterns, assuming a scene of spatial size 128×128, and three independent scenarios, where $s = 0.046$, $s = 0.093$, and $s = 0.1875$; according to Eq. (3.6), the number of recoverable spectral bands for each scenario is 6, 12 and 24, when β is assumed to be 1. The optimal patterns were compared against the Boolean and random coded apertures. As previously mentioned, the Boolean coded aperture patterns hold the condition $\sum_{q=1}^{Q} \mathbf{t}_{m,n,q} = 1 \;\forall\; m, n$; at the same time, the transmittance for each pattern is approximately $1/Q$. Unlike optimal coded apertures, Boolean coded apertures are randomly generated without considering the structure of the sensing matrix $\mathbf{A}$. The random coded apertures are generated such that the transmittance for every pattern is equal to 0.5, with no constraint about complementarity. Two criteria of comparison were implemented; the first one is the Peak Signal-to-Noise-Ratio ($PSNR$), defined in Eq. (2.32). The second criterion is the

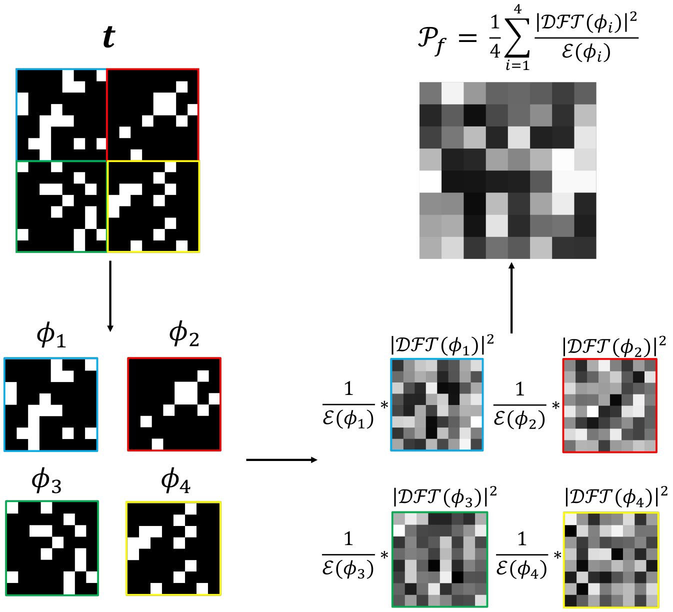

Figure 3.3: Illustration of the computation of the power spectral density $\mathcal{P}_f$, for a 16×16 pattern, using $R = 4$ subregions, ϕ_1, ϕ_2, ϕ_3 and ϕ_4. The coded aperture is first divided into the subregions; then, the normalized Power Spectral Density is individually calculated for the subregions. $\mathcal{P}_f$ is then defined as the average of the Power Spectral Densities.

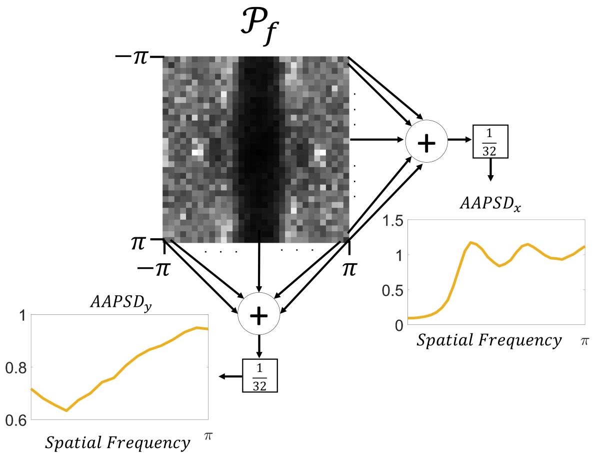

Figure 3.4: Illustration of the computation of *Axially Averaged Power Spectral Density* over the rows and the columns, for a given 32×32 $\mathcal{P}_f$.

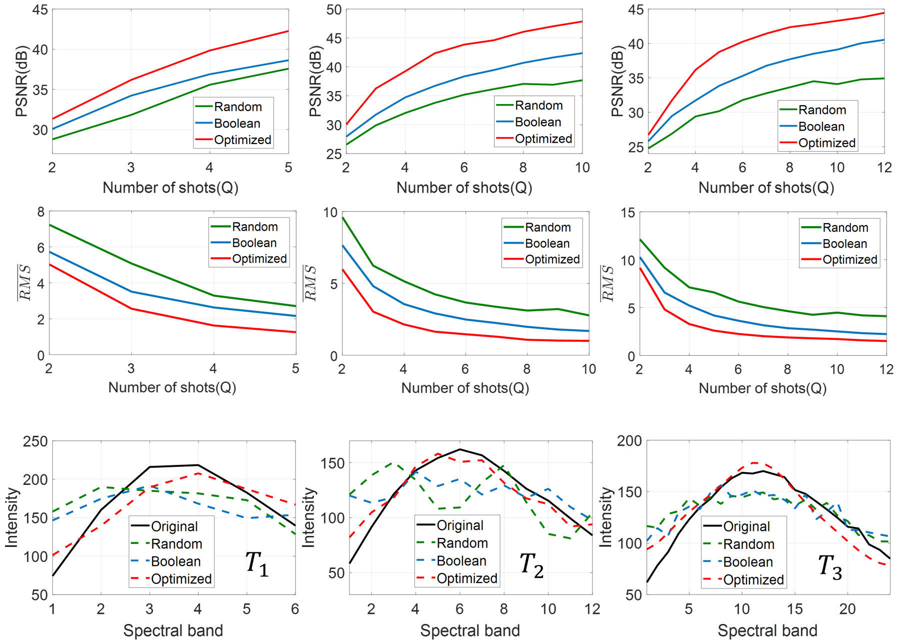

Figure 3.5: Top and Middle: $PSNR$ and $\overline{RMS}$ of the reconstructed scenes as a function of the number of captured snapshots for a $128 \times 128 \times 6$ scene and $s = 0.046$ (left), a $128 \times 128 \times 12$ scene and $s = 0.093$ (middle), and a $128 \times 128 \times 24$ scene and $s = 0.1875$ (right). Bottom: Original and reconstructed signatures for pixels T_1 (left), T_2 (middle) and T_3 (right) specified in Fig. 3.6.

Averaged Spectral Root Mean Squared Error ($\overline{RMS}$), defined as the average RMS of the spectral responses for every single pixel on the scene. As in the previous chapter, the GPSR algorithm [33] is implemented to iteratively solve Eq. (1.2), by tuning τ to obtain optimal reconstruction results. Figure 3.5 shows the reconstructed spectral signatures for different pixels and the $PSNR$ and $\overline{RMS}$ of the recovered scene as a function of the captured snapshots, while Figure 3.6 shows the RGB-mapping of the original and reconstructed scenes using the different pattern structures. As depicted, the optimized patterns overcome the performance of the random and Boolean coded apertures.

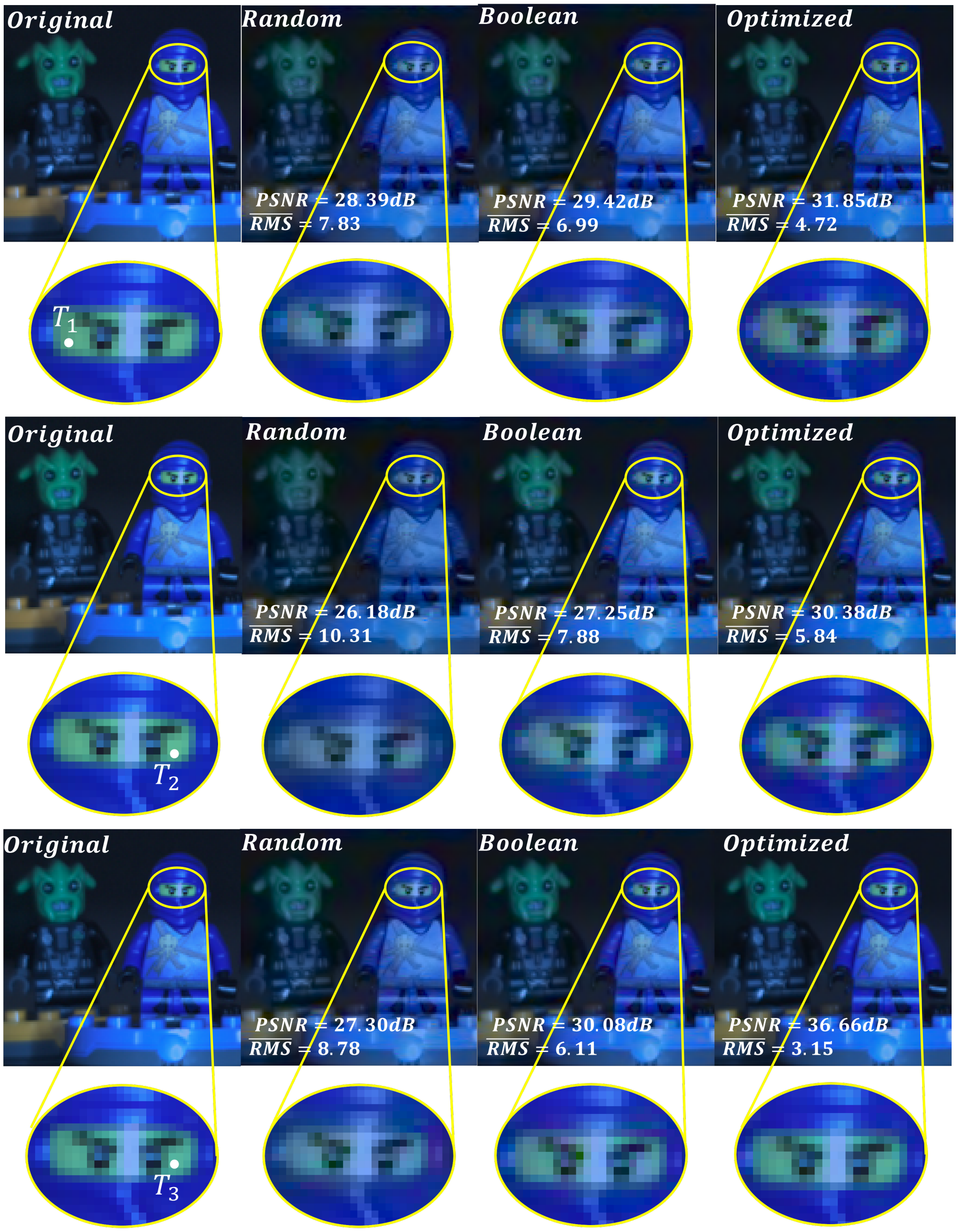

Figure 3.6: RGB-mapping of the original and reconstructed datacubes using random, Boolean and optimized patterns. Top: $s = 0.046$, two snapshots and a $128 \times 128 \times 6$ datacube. Middle: $s = 0.093$, two snapshots and a $128 \times 128 \times 12$ datacube. Bottom: $s = 0.1875$, four snapshots and a $128 \times 128 \times 24$ datacube.

3.6.1 Convergence and complexity of the algorithm

In order to test the convergence of the proposed algorithm, 100 independent executions were run for two different scenarios; $s = 0.093$ and eight snapshots, and $s = 0.1875$ and eight snapshots. The convergence of the objective function for the previously mentioned scenarios can be seen in Fig. 3.7. In both scenarios, the objective function is minimized as the number of iterations increases. In terms of computational complexity of the proposed algorithm, Eq. (3.17) entails a computational complexity of $\mathcal{O}(N^2L^2)$, while the calculation of $\hat{F}(q)$ (line 10 in Algorithm 1) has a computational complexity of $\mathcal{O}(2L^2)$, assuming that the set $\mathbf{I} = \mathbf{Find}(r_1, r_2)$ has an extension of $2L$. Also, lines 1, 3, 5, 7, 8, 12, 13, 16 and 17 require a simple operation, $\mathcal{O}(1)$. The computational complexity of Algorithm 1 can be seen in Table 3.1.

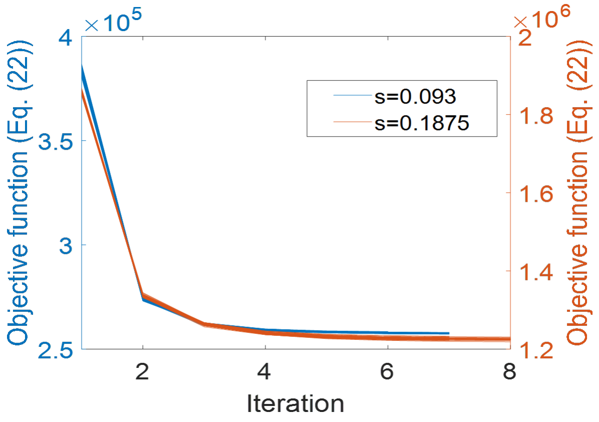

Figure 3.7: Convergence of the objective function, given by Eq. (3.17), for $s = 0.093$, eight snapshots, and a $128 \times 128 \times 12$ recovered datacube (blue lines), and for $s = 0.1875$, eight snapshots and a $128 \times 128 \times 24$ recovered datacube (red lines). The algorithm was run 100 times; the lines represent each execution. When $s = 0.093$, 73 out of the 100 executions reach seven iterations; 27 out of 100 executions reach six iterations. When $s = 0.1875$, 18 out of the 100 executions reach eight iterations; 82 out of 100 executions reach seven iterations.

Table 3.1: Computational complexity of Algorithm 1, in terms of the total number of iterations IT, number of captured snapshots Q, and size of the datacube $N \times N \times L$.

Lines	Operation	Complexity
2	Eq. (3.17)	$\mathcal{O}(N^2L^2)$
4 to 14	**for** loop over R, **for** loop over Q	$\mathcal{O}(N^2Q2L^2)$
15	Eq. (3.17)	$\mathcal{O}(N^2L^2)$
1 to 18	Full algorithm	$\mathcal{O}(2N^2L^2(1+Q)\text{IT})$

3.6.2 Noisy system

In order to test the robustness of the optimal codes against noise, the reconstructions were run for two different levels of Signal to Noise Ratio (SNR), 30dB and 50dB. Here, the noise is assumed to be Additive White Gaussian (AWGN). The results can be seen in Fig. 3.8. Notice that, when the system SNR is 30dB, the $PSNR$ of the reconstructed datacube starts to decrease compared to the noiseless scenario.

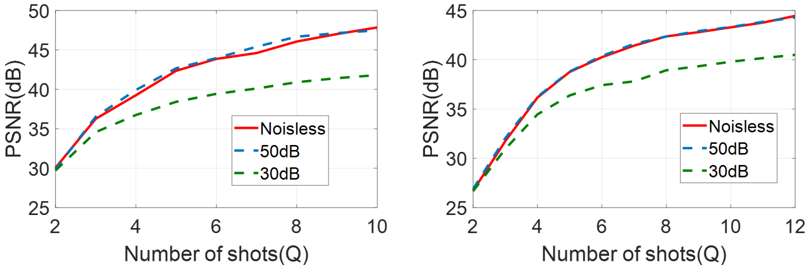

Figure 3.8: $PSNR$ of the reconstructed datacube for two levels of SNR, 30dB and 50dB. Left: $s = 0.093$ and the datacube has dimensions of $128 \times 128 \times 12$. Right: $s = 0.1875$ and the datacube has dimensions of $128 \times 128 \times 24$. In both cases, the obtained curves are compared to the noiseless scenario.

3.6.3 Performance for different hyperspectral scenes

In order to test the robustness of the optimal patterns, the reconstruction algorithm was run for two different hyperspectral scenes that belong to the Computer Vision Laboratory (CAVE) at Columbia University multispectral image database [61]. Figure 3.9 shows the RGB-mapping of the original and reconstructed $128 \times 128 \times 24$ datacubes for 3 captured snapshots using random, Boolean and the optimized patterns; here, $s = 0.1875$ in order to recover the 24 spectral bands, and assuming $\beta = 1$, in Eq. (3.6). A comparison of the $PSNR$ and $\overline{RMS}$ of the recovered scenes as a function of the number of captured snapshots can be seen in Fig. 3.10; the recovered spectral signatures for the pixels specified in Fig. 3.9, can be seen in Fig. 3.11.

3.6.4 Performance of the optimal patterns for bigger datacubes

The proposed algorithm was run to recover a $256 \times 256 \times 24$ beads datacube. This is done to evaluate the performance of the algorithm for bigger spectral scenes. According to Eq. (3.6), assuming $\beta = 1$, to recover 24 spectral bands, s must be equal to 0.093 if $N = 256$. Figure 3.12 shows the RGB-mapping of the original and reconstructed scenes, using random, Boolean, and optimized patterns; here, eight snapshots were captured. Notice that the reconstruction when using optimized codes shows better results in $PSNR$ and $\overline{RMS}$. The spectral RMS for all the pixels of the reconstructed scenes is also shown in Fig. 3.12; as it can be seen, the error in the reconstruction is considerably reduced when using the optimized codes.

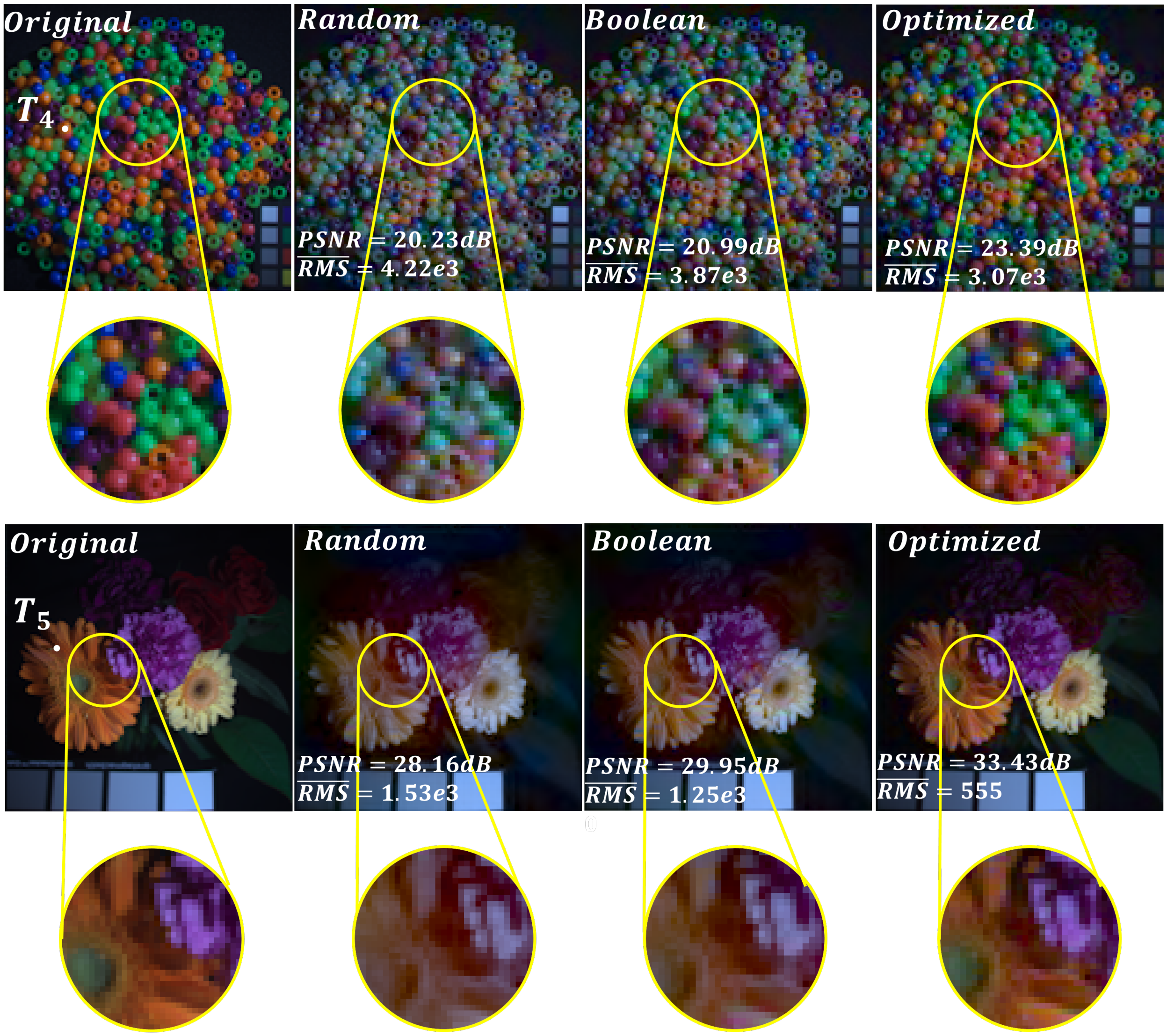

Figure 3.9: RGB-mapping of the 128 × 128 × 24 original and recovered datacubes, using random, Boolean and the optimized coded aperture patterns, three snapshots and $s = 0.1875$. Top: Beads datacube. Bottom: Flowers datacube.

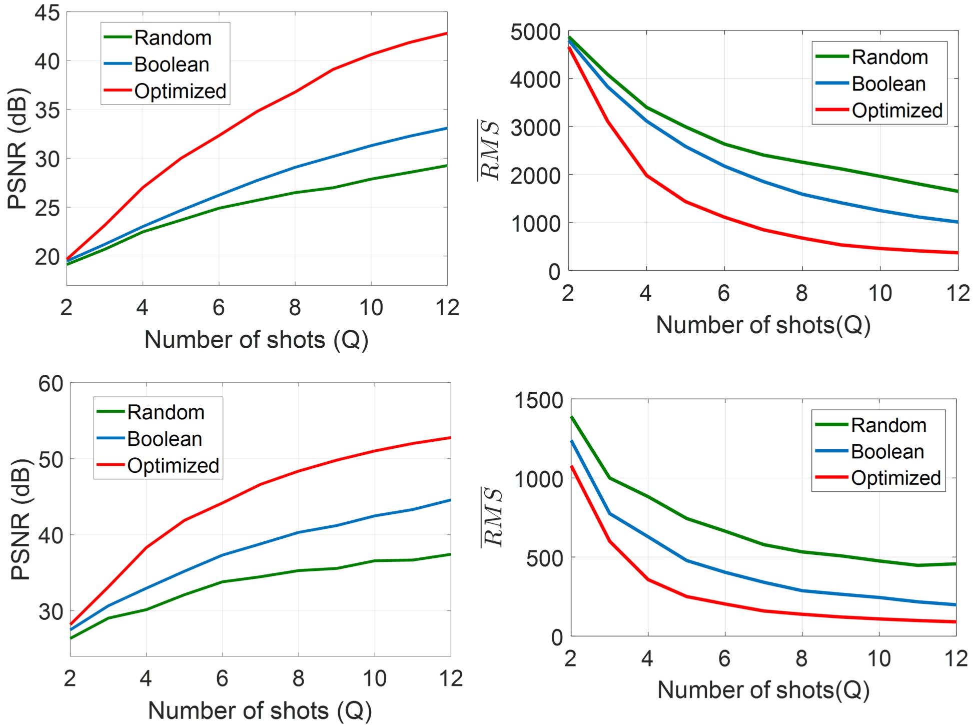

Figure 3.10: $PSNR$ and $\overline{RMS}$ of the recovered $128 \times 128 \times 24$ datacubes, with $s = 0.1875$, as a function of the captured snapshots, using random, Boolean and optimized patterns, for Beads datacube (top) and Flowers datacube (bottom).

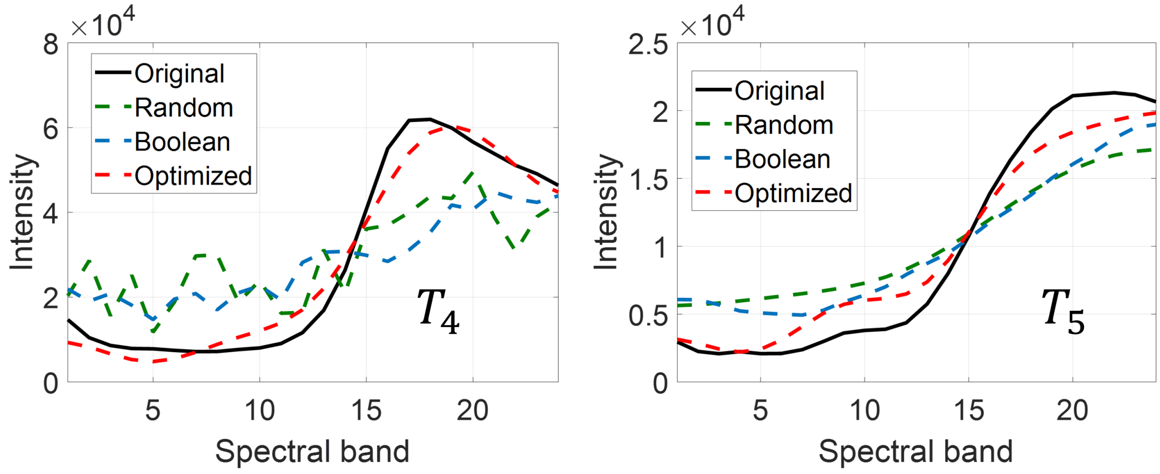

Figure 3.11: Recovered spectral signatures for pixels T_4 and T_5, specified in Fig. 3.9.

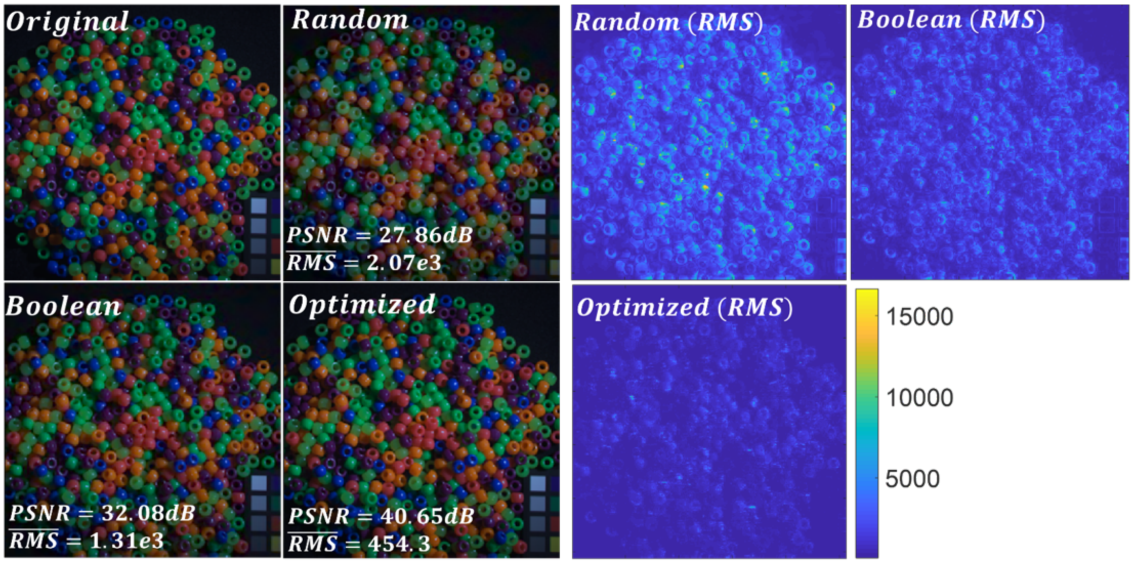

Figure 3.12: Left: RGB-mapping of the original and reconstructed 256 × 256 × 24 datacubes, using random, Boolean and the optimized codes. Here, $s = 0.093$ and eight snapshots were captured. Right: Spectral RMS for all the pixels of the reconstructed scenes.

3.6.5 Optimal codes structure

The structure of the Boolean and optimized patterns for the case of five snapshots and $s = 0.046$ can be seen in Fig. 3.13-top. Likewise, the Boolean and optimal patterns for 10 snapshots, $s = 0.093$ and $s = 0.1875$, can be seen in Fig. 3.13-bottom. Notice that, for a given row of the optimal patterns, the pixels that are set to one are widely spread; this minimizes the off-diagonal elements of $\mathbf{A}^T\mathbf{A}$; at the same time, some clusters of the pixels that are set to one are present in the optimal patterns; this is related to the condition of the on-diagonal elements of $\mathbf{A}^T\mathbf{A}$. In addition, the clusters become more noticeable as s increases. One explanation for this particular behavior is that as s increases, the number of elements in $\mathbf{H}$, that are different from 0 or 1, also increases; this forces the algorithm to cluster more elements in order to reach the condition $(\mathbf{A}^T\mathbf{A})_{b,b} = 1 \ \forall b \ \in [0, 1]$. As s decreases, on the other hand, the elements in $\mathbf{H}$ that are different from 0 or 1 become negligible. This forces the algorithm to fully focus on the minimization of the off-diagonals elements of $\mathbf{A}^T\mathbf{A}$. The *Axially Averaged Power Spectral Density* over the rows ($AAPSD_x$) and columns ($AAPSD_y$) of the optimal coded apertures for the three values of s can be seen in Fig. 3.14. Notice that for $AAPSD_x$ (see Fig. 3.14-left), when $s = 0.046$, all the spectral content is located on the high frequencies. As s increases, the spectral content can be found in low to middle frequencies. For $AAPSD_y$ (see Fig. 3.14-right), the spectral content of the optimal codes shows less power in the low frequencies. The spectral content of a regular Boolean pattern, on the other hand, is homogeneously spread all over the spectrum.

3.7 Conclusion

In this chapter, an optimization framework for the SSCSI coded aperture patterns is proposed and developed, based on the discrete measurements model previously proposed in chapter 2. The optimal coded aperture patterns must exhibit the following properties: first, the ON pixels must be separated as much as possible over the rows of the pattern. This condition minimizes the values of the off-diagonal elements of $\mathbf{A}^T\mathbf{A}$, where $\mathbf{A}$ is the sensing matrix of the system. Secondly, some clusters of the ON pixels over the rows must appear on the final optimal patterns. This condition ensures on-diagonal elements of $\mathbf{A}^T\mathbf{A}$ are close to 1. Third, the patterns must be complementary over the captured snapshots, where complementarity was defined as having just one pixel set to ON at each spatial position of the ensemble of the patterns. The number of clusters on the final optimal patterns depends on the parameter s having fewer clusters as s is closer to zero. Using the *Axially Averaged Power Spectral Density*, a spectral statistic first proposed here, it was found that the optimal coded apertures exhibit blue noise characteristics over columns. The spectral content of the optimal patterns over the rows depends on the parameter s, showing blue noise characteristics when s is close to zero, and green noise characteristics as s increases. The obtained optimal patterns were tested against the Boolean complementary and random patterns, showing better

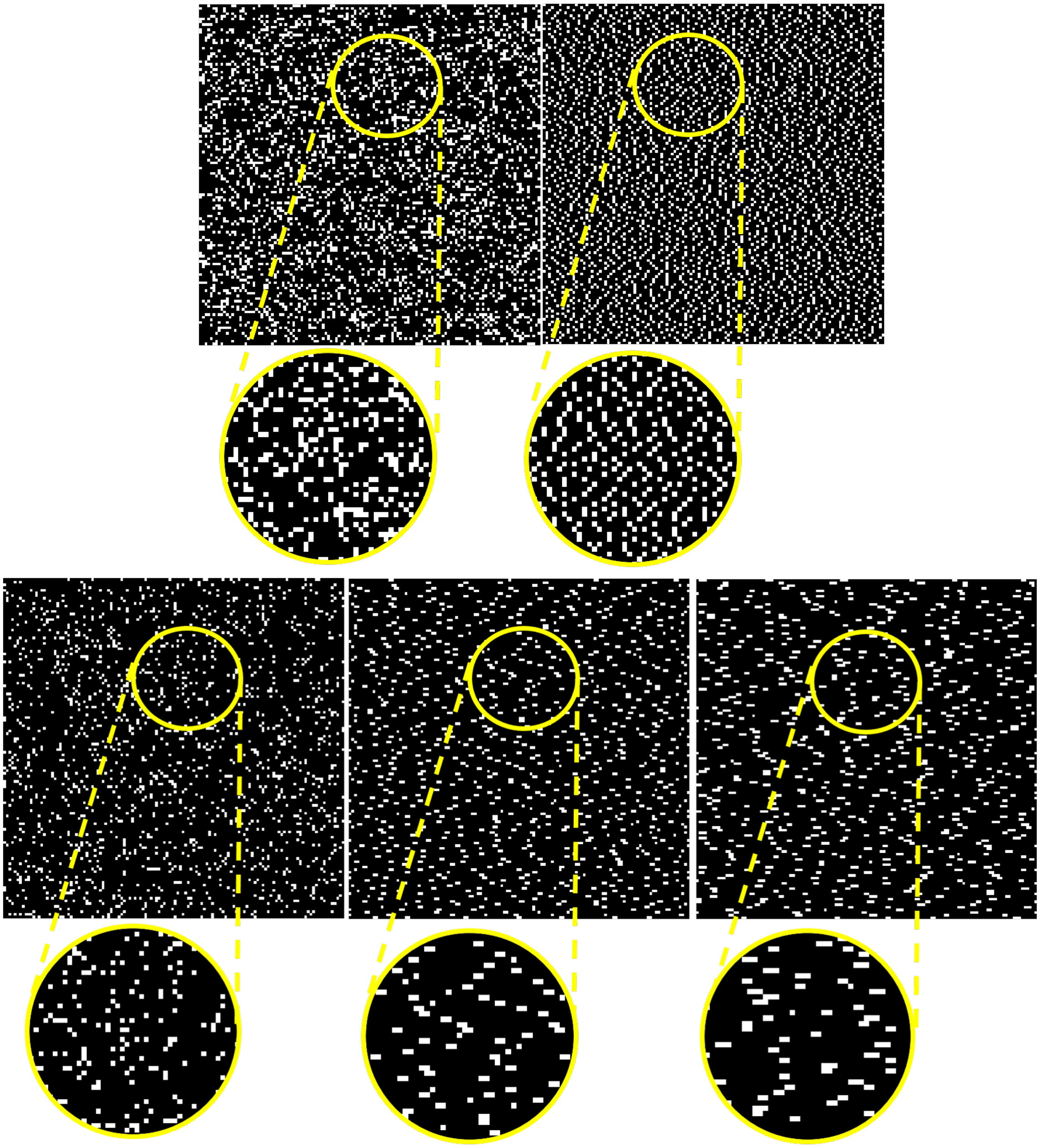

Figure 3.13: Top-left: Boolean pattern, for 5 snapshots. Top-right: Optimized pattern for 5 snapshots, when $s = 0.046$. Bottom-left: Boolean pattern, for 10 snapshots. Bottom-middle: Optimized pattern for 10 snapshots, when $s = 0.093$. Bottom-right: Optimized pattern for 10 snapshots, when $s = 0.1875$. Note how the ON pixels are separated as much as possible from each other over the rows in the zoomed version of the optimal coded aperture; at the same time, several clusters of ON pixels are present over the rows in the final patterns as s increases.

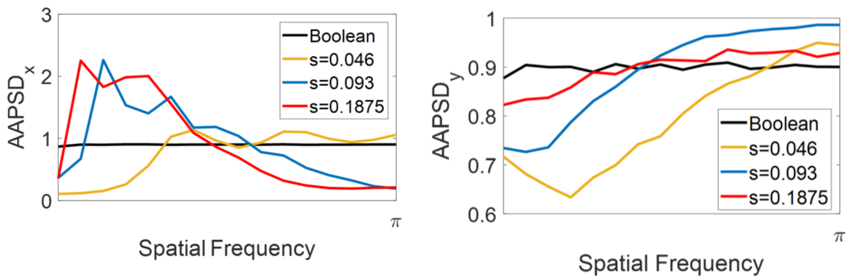

Figure 3.14: Axially Averaged Power Spectral Density of the boolean and optimized codes for $s = 0.046$, $s = 0.093$ and $s = 0.1875$. Here, the patterns were divided into $R = 16$ subregions. Left: $AAPSD$ over the rows. Notice that as s increases, the spectral content of the pattern can be found in low to middle frequencies. Right: $AAPSD$ over the columns. Here the spectral content of the optimal patterns shows less power at low frequencies.

performance than both of them for all the scenarios studied in this chapter (different hyperspectral scenes, different number of captured snapshots, and different values of s). The convergence and computational complexity of the proposed algorithm were analyzed.

Given that this is the first paper to successfully find optimal patterns in SSCSI, a conference paper in *Computational Optics, Sensing and Imaging*(Salazar et. al. [62]), and a journal paper in *IEEE transactions on Computational Imaging* (Salazar et. al.[63]), were published (see attached disclaimer for copyright issues). The work developed in this and the last chapter have been used as the basis for further research on SSCSI [64].

All the data and software for this chapter are available in [65].

Chapter 4

COMPRESSIVE X-RAYS COMPTON BACKSCATTERING IMAGING

4.1 Introduction

This research will now focus its efforts in adapting the coded aperture imaging principles previously defined for spectral imaging, to Compton Backscattering-based scanners. Compton backscattering imaging (CBI) is a technique that uses ionizing radiation in order to obtain information about the molecular composition of an object. It captures the scattered photons after radiation collides with the target; therefore, it is considered a single-sided technique that allows locating both the detector and target on the same side of the setup. CBI finds applications in organic material identification which are commonly encountered on explosives manufacturing, drugs, and contraband in general. This concept has been successfully applied to, for example, under-vehicle and airport luggage inspection and screening of buildings, among others [19, 20, 21]. Moreover, since the 9/11 terrorist attacks, CBI scanners have been applied to passenger screening at airports on a daily basis[66]. CBI can be broadly classified into two groups: pencil-beam and full-field illumination imaging [22]. The pencil-beam techniques rely on a highly collimated X-ray source impinging on the body and one or two scintillators used to capture the scattered photons. Most of the pencil-beam based scanners emulate the flying spot scheme [67], which is composed of a fan-beam collimator together with a rotating chopper wheel. In order to capture a full scene, the body under inspection must move horizontally. The functioning principle of the pencil-beam CBI scanning can be seen in Fig. 1.2. In full-field illumination techniques, the body is fully bathed by an X-ray cone-beam, and the pixels are registered parallel-wise. Naturally, recovering the wanted information is more challenging in this scenario. The captured photons are focused on a 2-dimensional sensor using either a coded aperture [24], X-rays optics (lobster eye) [23], spatially selective filters [68], or Fresnel plates [25].

The performance of a CBI architecture is dependent on different extrinsic and intrinsic factors. It has been shown that different levels of incident keV photons directly influence the final contrast of the image [69]. At the same time, the number of scattered photons turns out to be proportional to the thickness of the target, until it reaches the so-called *saturation thickness*, as widely explained by Hosamani et al. in [20]. On the other hand, the type and size of the scintillating material implemented in the conversion of high energy photons to visible light must be chosen in accordance with the expected scattered energy and intensity; commercial scanners implemented in security checkpoints are composed of sodium iodide(NaI) based crystals [66]; other scintillators in the literature have been designed to enhance the final optical transmittance of the system[70]. In this chapter, it is proposed a new architecture that allows the realization of CBI using structured light, under the theory behind compressive sensing [3]; the architecture is coined *compressive X-Rays Compton backscattering imager* (CXBI). CXBI emulates a simultaneous acquisition of different pixels by means of a coded aperture and a cone-beam X-rays source and separates such information through the solution of an optimization inverse problem. The functionality of CXBI is tested here with Montecarlo simulations, using the *Geant4 Application for Tomographic Emission (GATE)*, then an experimental test-bed is implemented in the laboratory.

The chapter is organized as follows: Section II raises the theoretical background of the Compton phenomenon; Section III introduces the new inverse problems approaches; Section IV describes in detail the proposed CXBI including the forward and discrete measurements model; Section V describes the GATE simulation conditions, while Section VI shows the results; Section VII makes an analysis of the needed dose in CXBI; Section VIII tests the reliability of CXBI for human inspection; Section IX shows the details of the conducted experiments; and finally, Section X closes this chapter with future steps and conclusions.

4.2 The Compton phenomenon

Compton scattering occurs when a photon and a charged particle (electron at rest) collide. This phenomenon is the predominant effect at low energies, together with the photoelectric effect. Assume a photon moves with velocity v_p and energy $E_p = hv_p$, where h is Plank's constant; if the photon hits an electron at rest with energy $E_o = m_oc^2$, with m_o being the electron mass at rest and c the speed of light, the photon will be scattered with energy E_s given by the following expression [71]:

$$E_s = \frac{E_p}{\frac{E_p}{E_o}\left(1 - cos(\theta)\right) + 1}, \tag{4.1}$$

where θ is the scattering angle of the photon. The expression in (4.1) is a direct result of the law of conservation of momentum. According to the literature, a photon is said to be backscattered if $\theta > \pi/2$ [69]. Notice that when $\theta = \pi$, the photon will be expelled with approximately 70% of the incident energy. The probability of Compton scattering events can be quantified using the differential scattering cross section, given by the Klein-Nishina formula [72]:

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2}\left(\frac{1}{1+\alpha\left(1 - cos(\theta)\right)}\right)^2\left(a + cos^2(\theta)\frac{\alpha^2(1-cos(\theta))^2}{1+\alpha(1-cos(\theta)}\right), \tag{4.2}$$

where $a = E_p/E_o$, $r_e = 2.818 \times 10^{-15}m$ is the classical electron radius, $d\sigma$ is the area involved in the scattering process, and $d\Omega$ is an infinitesimally solid angle element, subtended by the detector at the point of interest in the body. A graphical explanation of the scattering cross-section can be seen in Fig. 4.1. When collisions do not occur on the surface, but on a sub-surface level, the incoming as well as the scattered photons have to travel back and forth through the body, which attenuates its intensity through the Beer-Lambert law:

$$I = I_oe^{-\mu x}, \tag{4.3}$$

where I is the intensity after the beam has traveled a distance x on the material, I_o is the incident intensity, and μ is the linear attenuation coefficient coming from the

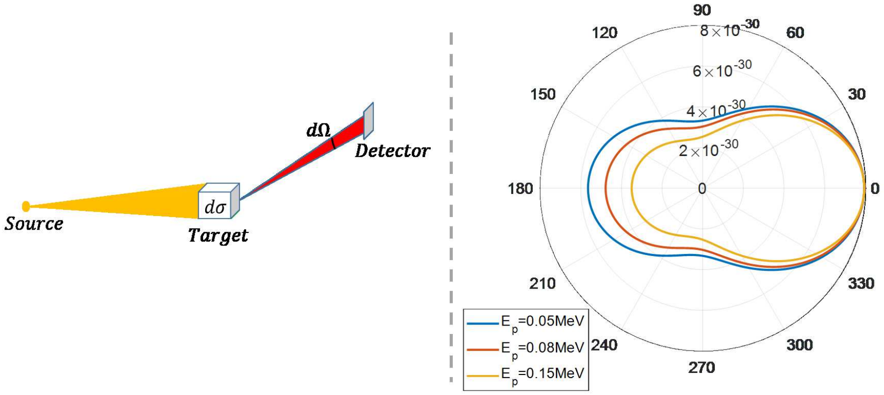

Figure 4.1: Left: Cross section involved in the Compton scattering phenomenon. Right: $\frac{d\sigma}{d\Omega}$ as a function of θ, for three different incident photon energies.

photoelectric effect and the Compton scattering. The possible attenuation due to *pair production* has been omitted, as this phenomenon occurs for high energy photons close to γ rays, which are out of the scope of this research.

4.3 New inverse problem approaches

In the last two chapters, the iterative algorithm GPSR was implemented due to its good performance in CSI. However, several new approaches have been proposed, and, with the advent of deep learning, neural networks start playing a fundamental role in the reconstructions. One of the key advances in inverse problems was the called *Alternating Direction Method of Multipliers* (ADMM) [73], which decouples Eq. (1.2) into two independent optimization problems as follows:

$$\begin{aligned}\mathbf{x}^{k+1} &= \min_{x} \frac{1}{2}||\mathbf{Hx}-\mathbf{y}||_2^2 + \frac{\rho}{2}||\mathbf{x}-\mathbf{z}^k+\mathbf{u}^k||_2^2\\ \mathbf{z}^{k+1} &= \mathcal{D}(\mathbf{x}^{k+1}+\mathbf{u}^k, \sigma^2=\lambda/\rho),\\ \mathbf{u}^{k+1} &= \mathbf{u}^k+\mathbf{x}^{k+1}-\mathbf{z}^{k+1},\end{aligned} \tag{4.4}$$

where $\mathbf{z}$ and $\mathbf{u}$ are intermediate variables coming from the reformulation of the problem using the Augmented Lagrangian Multipliers and $\mathcal{D}$ is a denoising operation; the solution of the first line in the last algorithm can be easily calculated given that this is a second order expression with defined global minimum, while the denoiser can be trained through deep learning techniques. This framework has been successfully adapted to CSI [74]. Other advances in inverse problems include the unfolding algorithms [75], deep convolutional autoenconders [76], recurrent neural networks [77] and generative adversarial networks [78]. In the context of Compton scattering, the concept of inverse problems has already been applied to recover a three-dimensional electron density map by means of a mono-energetic source, an array of spectral detectors, and assuming the signal dependent noise (Poisson) [79, 80]. Given that the forward model is not easily invertible, iterative methods have been used to find the electron density map. Moreover, the incompleteness of the data inherently arises from the energy resolution of the system and the limited angular view. To the best of our knowledge, this work is the first attempt to treat Compton Backscattering Imaging through inverse problems and coded illumination.

4.4 Compressive X-rays Compton backscattering imaging

As mentioned in the introduction of this chapter, CBI can be executed using either pencil-beam scanning or cone-beam illumination. Although the former delivers high quality results, the need of scanning the scene pixel-wise might represent a drawback. This dissertation proposes a novel scanning technique, where a random parallel acquisition of the scene pixels is done by means of coded illumination or structured light. A similar approach was already proposed under different contexts, such as the X-ray diffraction tomography [81, 82]. The proposed architecture can be seen in Fig. 4.2-right (the pencil-beam scanning architecture is shown in Fig. 4.2-left for comparison). A cone-beam illuminates a predefined portion of the coding pattern. After that, the structured light hits the body under inspection and the photons scattered by Compton scattering are registered by two scintillating plates. The cone beam origin is assumed to have a small focal spot, so that the penumbra effect is small [83]. As for the energy characteristics of the source, this paper tests CXBI under a poly-energetic X-rays source, given that mono-energetic sources are too costly for real applications.

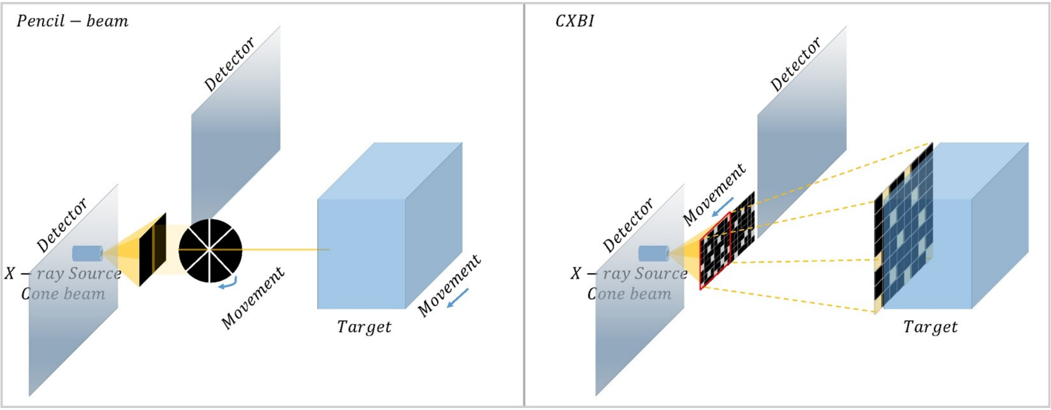

Figure 4.2: Left: Flying spot architecture. An X-rays cone-beam source passes through a fan beam collimator and posteriorly through a chopper wheel which acts as rotating pencil-beam collimator; this allows the body to be vertically scanned. In order to fully scan the body, it must translate horizontally. Right: The proposed compressive Compton X-rays backscattering imaging (CXBI). The structured light arrives to the body under inspection, conducting a random sampling over the field of view, while the coded aperture continuously moves. The system can also be conceived as a static coded aperture with a moving body, which is more practical.

A well detailed experimental study of CBI under poly-energetic sources is developed in [70]. After the structured light related to the pattern hits the body and the photons are captured, the pattern moves one column so the coded aperture in the field of view (red square) changes, and the captured pixels are overall different. The system can also be conceived as an static coded aperture with a moving body. This sensing principle was already implemented in the visible light regime, through the well known single pixel camera [76, 77, 78, 84].

4.4.1 Discrete model

4.4.1.1 Mono-energetic assumption

Let $B_{i,j}$ be the photon flux (number of photons per area per time) impinging on the region or pixel (i, j) in the body under inspection. The number of scattered photons by action of the Compton phenomenon can be calculated as follows:

$$P_{in} = \left(B_{i,j}N_1^2 t\right) e^{-\mu x} \frac{d\sigma_T}{d\Omega}, \tag{4.5}$$

where $N_1 \times N_1$ is the size of the pixel (i,j), t is the exposure time, x is the penetration depth of the photons inside the body, $e^{-\mu x}$ is the attenuation experienced by the photon flux in accordance to the Beer-Lambert law, and $\frac{d\sigma_T}{d\Omega}$ is the total cross-section involved in the scattering process, which is defined as:

$$\frac{d\sigma_T}{d\Omega} = N_1^2 dx \frac{d\sigma}{d\Omega} \mathcal{P} Z, \tag{4.6}$$

where $N_1 \times N_1 \times dx$ are the dimensions of the region where the Compton scattering events occur, Z is the effective atomic number, $\mathcal{P}$ is the number of atoms per unit volume, while $\frac{d\sigma}{d\Omega}$ is given by Eq. (4.2). Notice that $\mathcal{P}Z$ is an indicator of the electron density of the material involved in the interaction, a quantity that is intended to be recovered in Compton Scattering Tomography [85]. As expected, $\frac{d\sigma_T}{d\Omega}$ is inherently material dependent. The total number of scattered photons can be defined as:

$$P_s = \int_x \left(B_{i,j}N_1^2 t\right) e^{-\mu x} N_1^2 \frac{d\sigma}{d\Omega} \mathcal{P} Z dx. \tag{4.7}$$

The variable x must be integrated from zero, where the body initiates, up to a depth where 95% of the Compton scattering events occur [66]. Once the photons are scattered, they travel back towards the detector, passing again through the body and following a path of length ℓ_s with linear attenuation coefficient of μ_s; therefore, the intensity is attenuated by $e^{-\mu_s \ell_s}$. The variable l_s can be written as a function of the photon scattered angle θ as $\ell_s = \frac{x}{cos(\pi-\theta)}$. The number of captured photons by the two plates can be written as follows:

$$\mathbf{F}_{i,j} = \int_\psi \int_{\omega_{i,j}}^{\pi} \int_x \left(B_{i,j}N_1^2 t\right) e^{-\mu x} N_1^2 \frac{d\sigma}{d\Omega} \mathcal{P} Z e^{-\mu_s \frac{x}{\cos(\pi-\theta)}} dx d\theta d\psi, \tag{4.8}$$

where $\omega_{i,j}$ is the minimum angle of the scattered photons captured by the detectors which depends on the distance between the body under inspection and the detectors

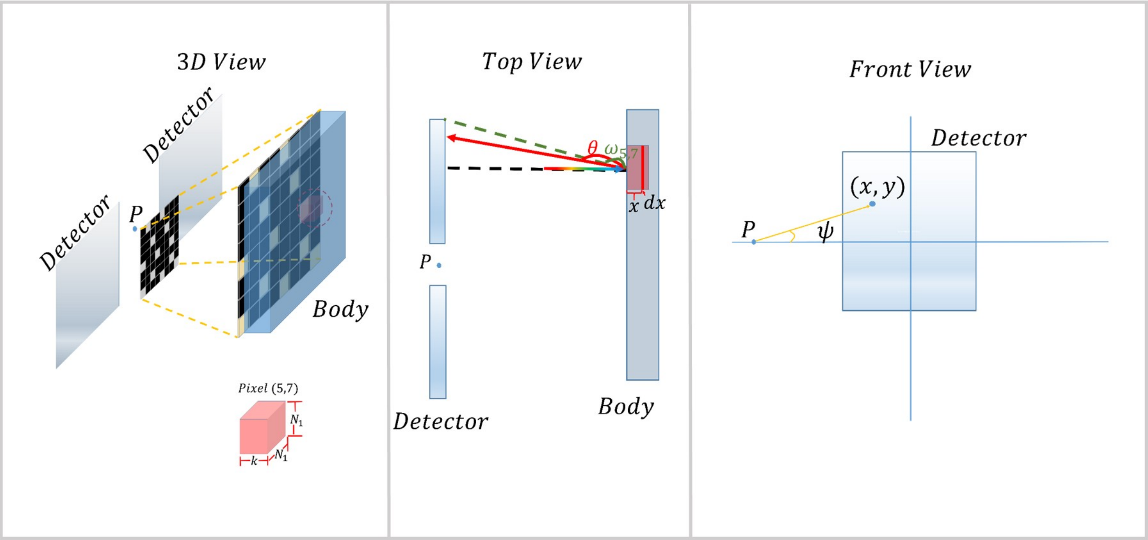

Figure 4.3: Graphical description of the discretization model. Left: 3D view of the CXBI and the scattering region of pixel (5,7). The region is a cube with dimensions $N_1 \times N_1 \times k$, where k is the depth up to which 95% of the single scattering events occur. Middle: Top view of the CXBI; the photon (multi-color arrow) arrives and it is scattered with angle θ (red arrow). The minimum scattering angle is $\omega_{5,7}$ (green dashed line) and the maximum scattered angle is π. Right: Front view of one of the detectors in CXBI. ψ is the angle between a middle point between the detectors, P, and a spatial position (x,y) on such detectors.

and on the dimensions of the scintillating plates. The variable ψ accounts for the different positions of the photons once they hit the detectors (see Fig. 4.3 for a better understanding of the mentioned variables). A similar analysis for the case of the pencil-beam scanner was developed by Cao [66].

4.4.1.2 Poly-energetic assumption

To extend Eq. (4.8) to the poly-energetic scenario, assume that the incoming energy $B_{i,j}^{E_p}$ is the photon flux per incident energy (or energy bin) at the region (i, j). The number of scattered photons captured by the detectors can be written as follows:

$$\mathbf{F}_{i,j} = \sum_{E_p} \tau(E_p) \int_{\psi} \int_{\omega_{i,j}}^{\pi} \int_{x} \left(B_{i,j}^{E_p} N_1^2 t \right) e^{-\mu x} N_1^2 \frac{d\sigma}{d\Omega} \mathcal{P} Z e^{-\mu_s \frac{x}{\cos(\pi-\theta)}} dx\, d\theta\, d\psi, \tag{4.9}$$

where $\tau(E_p)$ is the estimated contribution of each energy bin, in percentage, to the final number of captured backscattered photons. That is how, for example, for a $50kVp$ source filtered with a 1mm Aluminum window, 17% of the scattered photons belong to the energy region $0-10keV$, 41.3% to $10-20keV$, 31% to $30-40keV$ and 10.7% to the region $40-50keV$ [86]. Cao in [66] developed a similar analysis that applies for the pencil-beam scanner. Notice that, although not explicit, several variables of the argument in the last integral depend on E_p, such as the scattering angle (see Eq. (4.1)), the followed path by the scattered photons l_s, and consequently, its linear attenuation coefficient μ_s.

4.4.2 Forward model

The CXBI sensing process will now be described by using a matrix-vector notation. Let $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_{N+Q-1}]$ be $N \times N + Q - 1$ matrix containing the pattern to be used during the capturing process in CXBI, where $\mathbf{t}_m$ is the m^{th} column. Here, $N \times N$ is the size of the projected pattern in every shot, and Q is the number of captured measurements. The pattern corresponding to the q^{th} snapshot can be isolated as

$$\mathbf{C}_q = \mathbf{T} \cdot \mathbf{B}_q, \tag{4.10}$$

where $q \in [0, Q-1]$ and $\mathbf{B}_q$ is defined as follows:

$$\mathbf{B}_q = \left[\mathbf{0}_{q \times N}; \mathbf{I}_{N \times N}; \mathbf{0}_{Q-1-q \times N}\right], \tag{4.11}$$

where $\mathbf{0}_{q \times N}$ is a matrix full of zeros with dimensions $q \times N$, $\mathbf{I}_{N \times N}$ is the identity matrix and $\mathbf{0}_{Q-1-q \times N}$ is a matrix full of zeros with dimensions $Q-1-q \times N$. The forward matrix $\mathbf{A}$, with dimensions $Q \times N^2$ can be defined as:

$$\mathbf{A} = \left[\vec{\mathbf{C}}_0, \vec{\mathbf{C}}_1, \cdots \vec{\mathbf{C}}_{Q-1}\right]^T, \tag{4.12}$$

where $\vec{\mathbf{C}}_q$ is the vector-wise version of $\mathbf{C}_q$. The captured information is therefore given by the following equation

$$\mathbf{y} = \mathbf{A} \cdot \vec{\mathbf{F}}, \tag{4.13}$$

where $\vec{\mathbf{F}}$ is the vector-wise version of $\mathbf{F}$ given in Eq. (4.9) with length N^2, and $\mathbf{y}$ is a column vector of length Q. Note that the formulated discretization model does not consider Compton multiple scattering processes, Rayleigh phenomenon and unwanted scattering processes coming from air and other bodies. Therefore, the real measurements are corrupted by noise

$$\mathbf{y} = \mathbf{A} \cdot \vec{\mathbf{F}} + \epsilon. \tag{4.14}$$

In this chapter, the source of noise is assumed to be independent from the acquired signal; this will be properly justified in the simulations. A graphical description of the ensemble of the matrix $\mathbf{A}$ can be seen in Fig. 4.4-top. An example of the sensing matrix for 10 captured snapshots and a 16×16 scene can be seen in Fig. 4.4-bottom. In this case, $\mathbf{A}$ has dimensions of 10×256. Notice in the zoomed portion how patterns relate to each other by a movement that represents the physical translation of the mask, or equivalently, the body.

4.5 GATE experiments

The proposed CXBI was mounted and tested using Geant4 Application for Tomographic Emissions (GATE), which is a user-friendly software that allows simulation of high energy physics experiments with conditions close to real life.

4.5.1 X-rays source

The Poly-energetic source was simulated using the Matlab tool SPEKTR [87]. The X-ray source was powered by a 120kV source and it was inherently filtrated by a 1.5mm aluminum window to be in accordance with standard of commercial scanners [66, 88]. The source beam expansion is 17.5 degrees and it is located at position

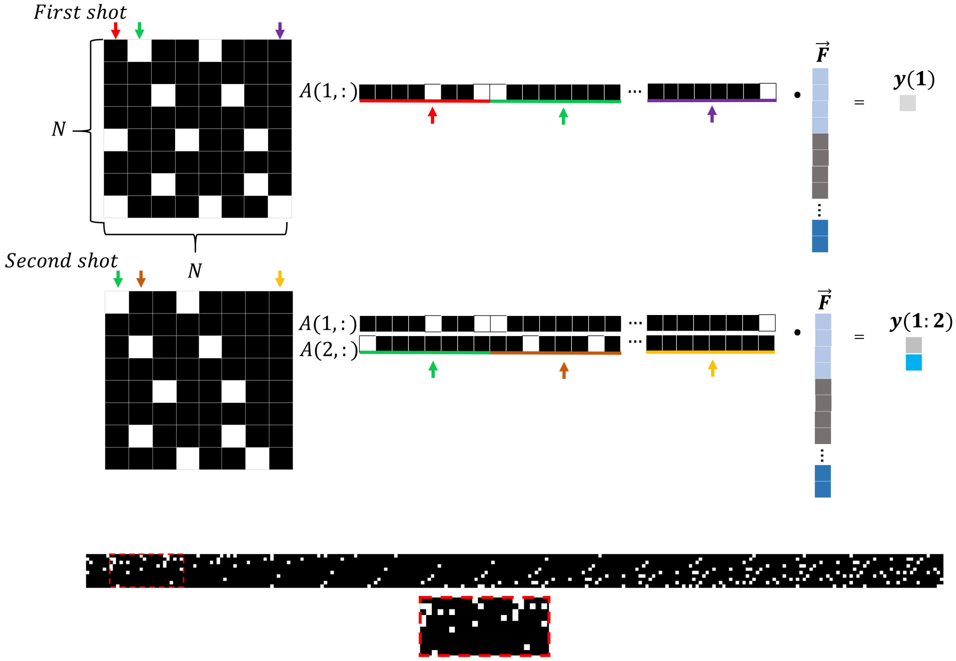

Figure 4.4: Top: CXBI forward model for two captured snapshots. Each row in $\mathbf{A}$ represents the row-wise pattern for a captured snapshot. Bottom: Sensing matrix for 10 captured snapshots and a 16 × 16 scene. The zoomed portion allows to see the movement of the rows patterns which is related to the physical translation of the mask.

(0,0,-15) cm. The activity of the source is set to $6.4 \times 10^8 Bq$, where $1Bq$ (Becquerel) is defined as the generation of one photon per second. By multiplying 6.4×10^8 by 1.6×10^{-19} (charge of electron) and divide it by 0.02 (percentage of electrons in an electron beam that generate X-rays [89, 90]), one obtains that the equivalent current is $5.12nA$.

4.5.2 Coded aperture

The coded aperture is conceived as a 32 × 32 random mask, with pitch size of $2mm \times 2mm$ and depth of $2mm$. The material of the coded aperture is Tungsten (W) because this element possesses a high atomic number and is suitable to project the pattern. The coded aperture is surrounded by a shielding wall (collimator) in order to avoid unwanted photons from the source colliding with the target. The coded aperture is positioned at the origin (0,0,0).

4.5.3 Detectors

Two detectors were placed between the target and the coded aperture in order to capture the backscattered photons after they collide on the body under inspection. The dimensions of the detectors are $30cm \times 50cm \times 5cm$ and the scintillating material is sodium iodide (NaI). The area was chosen so that the setup can be replicated in the laboratory. The thickness was chosen in order to have a high conversion ratio [70]. The detectors were located at (25,0,20.5)cm and (-25,0,20.5)cm. This location was set so that the plates are as close as possible to the target without being directly affected by the radiation of the source. In GATE, the output parameters in the simulations are known as *singles*. One *single* represents the aggregation of photon hits occurring at the detector layer. This aggregation emulates the integration in electronic hardware. For more details regarding the hits adding policies and more, please refer to GATE documentation [91].

4.5.4 Target

The body under inspection is placed such that its front face is located at (0,0,50)cm. The $2mm \times 2mm$ pixel of the mask is seen as a $0.867cm \times 0.867cm$ square at the body. This means that the spatial size of the recovered scene is limited to $27.7cm \times 27.7cm$, under proper conditions.

4.5.5 Capturing and reconstruction process

The capturing process is done snapshot by snapshot, and each snapshot takes $15.625ms$. The coded aperture translates with a uniform speed of $128mm/s$, such that each $15.625ms$ a new column enters and other exits the field of view. The number of singles per snapshot is extracted using ROOT and C++ to then concatenate the measurements vector $\mathbf{y}$. Five hundred and twelve shots were captured, which corresponds to 50% of the total number of pixels. There is no need to add artificial noise to $\mathbf{y}$ given that GATE emulates real experimental conditions. Only the photons arriving at the detector with an energy bigger than $10keV$ are taken into account when assembling $\mathbf{y}$. The matrix $\mathbf{A}$ is constructed according to the patterns in the snapshots, as depicted in Fig. 4.4. Finally the reconstruction algorithm is run to find $\mathbf{F}$.

4.5.6 Air and surrounding scattering reduction

Given that the photons interact with air and other bodies in the surroundings, unwanted scattering effects add noise to the captured data. In order to reduce its effects on the recovered image, the capturing process is run without target, such that the resultant measurements are subtracted from the ones with body.

4.5.7 Ground-truth capturing process

CXBI is a new architecture first proposed here; therefore, there is not available image database that fits the current needs of this research; hence the ground-truth images must be self-captured using GATE. The ground-truth was captured by moving a pencil-beam source pixel by pixel over the scene and then registering the backscattered photons. The pencil-beam was emulated by reducing the beam expansion of the same X-ray source to 0.4 degrees and it was located at $(x, y, 25)cm$, where x, y represents the spatial location of a given pixel over the target. The activity of the source was reduced to 5000 photons per second. The number 5000 was estimated as the mean number of particles that pass through a $2mm \times 2mm$ window during $15.625ms$ when the activity is set to $6.4 \times 10^8 Bq$.

4.6 Simulation results

Three different targets were implemented in the simulations. The first target was composed of four cubes of size $7cm \times 7cm \times 1cm$, equidistantly located over the origin. Two of them are composed of caffeine ($C_8H_{10}N_4O_2$, density $1.23g/cm^3$), and the remaining ones are composed of HMX-Octogen ($C_8H_8N_8O_8$, density $1.91g/cm^3$); HMX-Octogen is mainly used for plastic explosives [69]. The second target is composed of the letters U and D, both made of Aluminum (Z=13, density $2.7g/cm^3$). The last target was a human hand with a gold ring (Z=79, density $19.3g/cm^3$) located on the center finger. The density and composition of the hand is assumed to be close to that of the water. Figures 4.5, 4.6, and 4.7 show the ground-truth image, and the reconstructed images for code transmittances of 1%, 10%, and 50% per snapshot respectively; the code transmittance is defined as the ratio between the non-blocking pixels and the total number of pixels in the mask; in each figure, compression ratios of 12.5% and 25% are used, where the compression ratio is defined as the ratio between the number of captured measurements (length of $\mathbf{y}$) and the number of pixels on the scene (N^2). The chosen reconstruction algorithm was ADMM (Eq. (4.4)) with BM3D [92] and FFDNET (pre-trained) [93] in the denoising stage with variables initialized at zero. In order to decrease the correlation between adjacent masks, the patterns were chosen equidistantly over the originally-captured 512 shots. That is, when using a length of 12.5%, patterns were chosen every other 4 (50/12.5). The standard deviation of the noise in BM3D and FFDNET is chosen such that the best reconstruction is reached (noise level is unknown). Two criteria of comparison with the ground-truth are used, the structural similarity index (SSIM) and the Peak Signal-to-noise ratio (PSNR), defined in Eq. (2.32). The obtained SSIM and PSNR for the different scenarios are consigned in Fig. 4.5 (bottom-right and top-right respectively).

4.6.1 Increasing the resolution

Two main factors alter the resolution of the recovered images using CXBI. The first is the radiation of the X-ray source; a considerable decrease in the source activity makes the percentage of scattered photons too low to be distinguishable. Figure 4.8 shows the reconstructed UD target using $1.28nA$, $2.56nA$, and $5.12nA$, with a coded transmittance of 10% and a compression ratio of 12.5% and 25%. Notice that when the radiation is not enough, details in the reconstruction are missed; this is reflected in the SSIM and PSNR. The other factor that affects the final resolution is the pixel pitch size of the coded aperture; a reduction of the mask pitch size allows for the recovery of a scene with finer spatial details, as long as the radiation per pixel remains the same. Figure 4.9 shows the reconstruction of the UD target when decreasing the pixel pitch

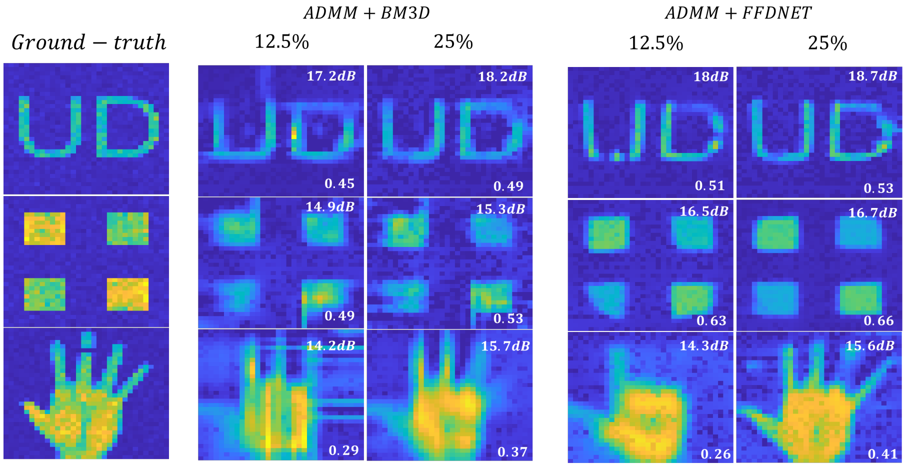

Figure 4.5: Ground-truth and reconstructed 32×32 images using 12.5% and 25% as compression ratio, and a code transmittance of 1% per snapshot. The ADMM was implemented using BM3D and FFDNET in the denoising stage. The SSIM and PSNR for every scenario can be observed in the bottom-right and top-right of the figures respectively.

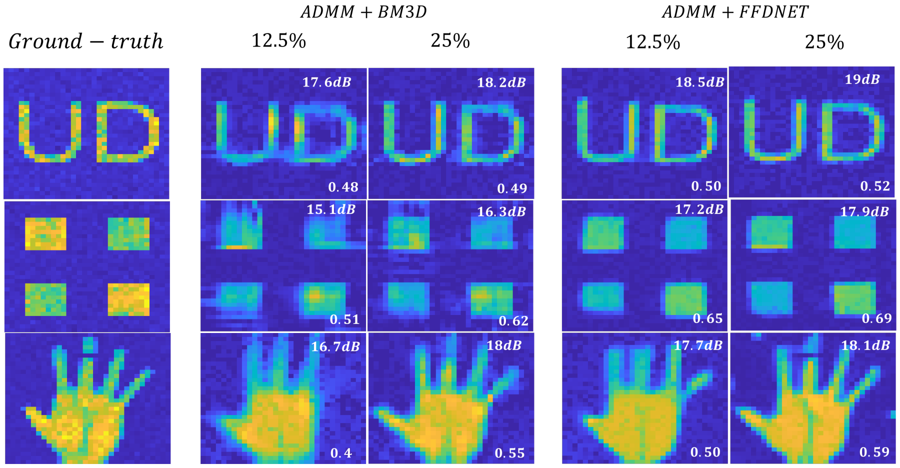

Figure 4.6: Ground-truth and reconstructed 32×32 images using 12.5% and 25% as compression ratio, and a code transmittance of 10% per snapshot. The ADMM was implemented using BM3D and FFDNET in the denoising stage. The SSIM and PSNR for every scenario can be observed in the bottom-right and top-right of the figures respectively.

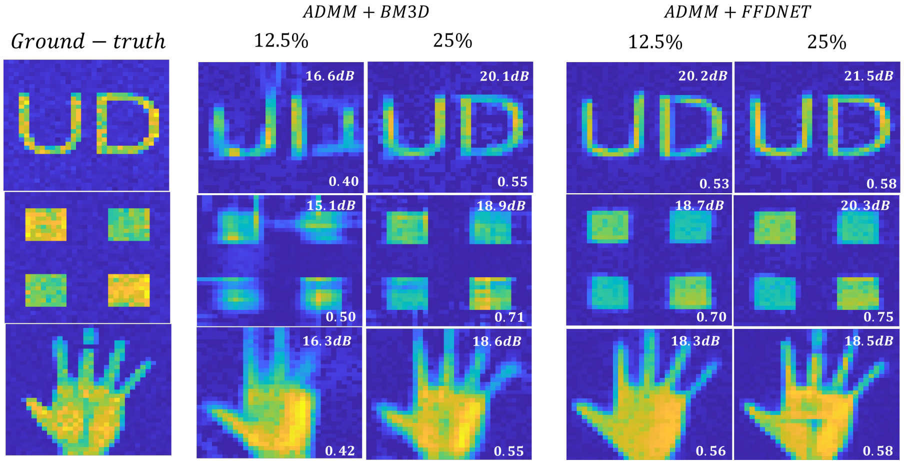

Figure 4.7: Ground-truth and reconstructed 32×32 images using 12.5% and 25% as compression ratio, and a code transmittance of 50% per snapshot. The ADMM was implemented using BM3D and FFDNET in the denoising stage. The SSIM and PSNR for every scenario can be observed in the bottom-right and top-right of the figures respectively.

size to $1mm$, using a current of $5.12nA$, and the same images using a pixel pitch size of $2mm$ and a current of $5.12nA$; as depicted, finer details can be seen when using a smaller pitch size; however, due to the fact that the same radiation distributes over a bigger quantity of pixels, the final obtained contrast abruptly decreases.

4.6.2 Noise Analysis

In CBI, the capturing process is usually assumed to be corrupted by signal dependent Poisson noise that models the inter-arrival time among photons. This is true when the photon arriving at the sensor in a given time is less than 100 [94]. The nature of the proposed CXBI, where there exists a parallel acquisition of pixels, allows one to safely assume that the level of Poisson noise is negligible. To test that, the mean number of photons arriving at the detectors is measured and consigned in table 4.1 for the simulation results in Figs. 4.5, 4.6, and 4.7. Notice that the number of photons arriving at the detector per captured snapshot is sufficient to safely neglect Poisson noise, even for the smallest transmittance per snapshot.

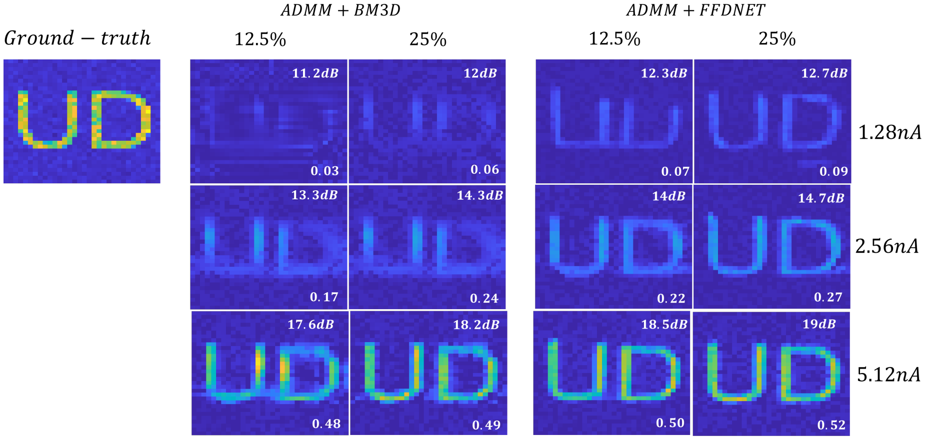

Figure 4.8: Ground-truth and reconstructed 32×32 images using 12.5% and 25% as compression ratio, for three different levels of source radiation: $1.28nA$, $2.56nA$, and $5.12nA$. The two denoisers (BM3D and FFDNET) were implemented in the reconstructions. The code transmittance is fixed to 10%.

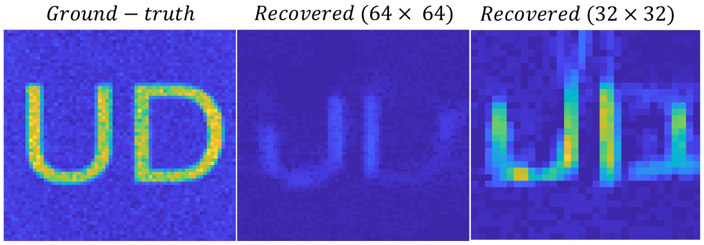

Figure 4.9: Ground-truth and reconstructed 64×64 images using 12.5% as compression ratio (left and middle) and its equivalent reconstructed 32×32 images, using an activity of $5.12nA$. Notice that, although fine details might be recovered when decreasing the coded aperture pitch size, due to the fact that the same radiation distributes over a bigger quantity of pixels, the final obtained contrast abruptly decreases.

Table 4.1: Mean number of photons arriving at the detectors for one single measurement for the three different targets in Figs. 4.5, 4.6, and 4.7.

	Mean number of photons arriving at the detectors		
Transmittance per shot	1%	10%	50%
UD letters	1.3×10^3	3.2×10^3	1.27×10^4
Squares	1.4×10^3	4.5×10^3	1.9×10^4
Human hand	1.9×10^3	8.7×10^3	4×10^4

4.7 Dose analysis

One of the key advantages of CBI is the level of radiation dose impinging on the body under inspection. A short time acquisition makes the radiation exposure negligible [68]. For example, the ZBV (Z Back-scanner Van) [21], which is a mobile system that scans vehicles using a pencil-beam collimated mono-energetic source of 225keV, requires a radiation dose of 0.07 μSv ($Sv \rightarrow Siever$) per inspection [68]. The absorbed radiation for a given pixel (i, j) is defined as follows:

$$dose(E_p) = \sum_{E_p} \frac{E_p B_{i,j}^{E_p} N_1^2 t \mu_{en,m}}{N_1^2}, \tag{4.15}$$

where $E_p B_{i,j}^{E_p} N_1^2 t$ is the total impinging energy in Joules into the pixel (i, j), and $\mu_{en,m}$ is the mass energy absorption coefficient given in cm^2/kg. The final units of the absorbed radiation are $Grays(Gy)$ ($Gy = \frac{J}{kg}$). Table 4.2 shows the estimated absorbed dose per pixel in the body in one CXBI shot, in a muscle-skeleton phantom. The coded aperture pitch size varies from $0.5mm$ to $4mm$, while the activity of the source remains constant at $5.12nA$. The curve representing $\mu_{en,m}$ for a muscle skeleton phantom can be seen in Fig. 4.10-left [95]. The number of photons per pixel in the body, and the area of such pixel are also consigned in Table 4.2.

Table 4.2: Absorbed dose per pixel in one snapshot, number of photons arriving at a pixel in one snapshot, and area of the pixel, for different coded aperture pitch sizes, and a source activity of $5.12nA$. The mass absorption coefficient values can be seen in Fig. 4.10-left

Pixel pitch size (mm)	Absorbed dose (nGy)	Number of photons	Area (cm^2)
0.5	30.9	401	0.05
1	10.9	1480	0.19
2	7.1	5286	0.75
4	4.4	23187	3

Figure 4.10-right shows the original and reconstructed human hand, using three different levels of absorbed dose, $0.035 \times 10^{-4}Gy$, $0.34 \times 10^{-4}Gy$, and $1.70 \times 10^{-4}Gy$ (top of each figure), and its respective SSIM (bottom-left), and PSNR (bottom-right); the compression ratio remains fixed at 12.5%, and the dose is controlled with the transmittance per captured snapshot. Notice that, in all cases, the absorbed dose is less than the one needed to capture the ground truth. The main reason behind this is the fact that a highly collimated pencil-beam is used to run the pixel-by-pixel scanning (see section 4.5.7).

4.8 CXBI for human inspection

The proposed architecture can potentially be implemented to improve the speed of scanning in critical checkpoints such as airports and homeland security borders. In order to prove that, a human male of 5 feet tall with a suspected gun on him is scanned using the conventional pencil-beam with activity of $3000Bq$, to then synthetically generate the CXBI measurements vector. The scanner dimensions are modified to $30cm \times 176cm \times 5cm$. To test the CXBI in a noiser environment, artificial Gaussian noise with variance of $\sigma^2 = 10000$ is added. The body composition and density are assumed to be close to that of the water ($\rho = 1g/cm^3$, 11.2% $H, 88.8\%\ O$); as for the

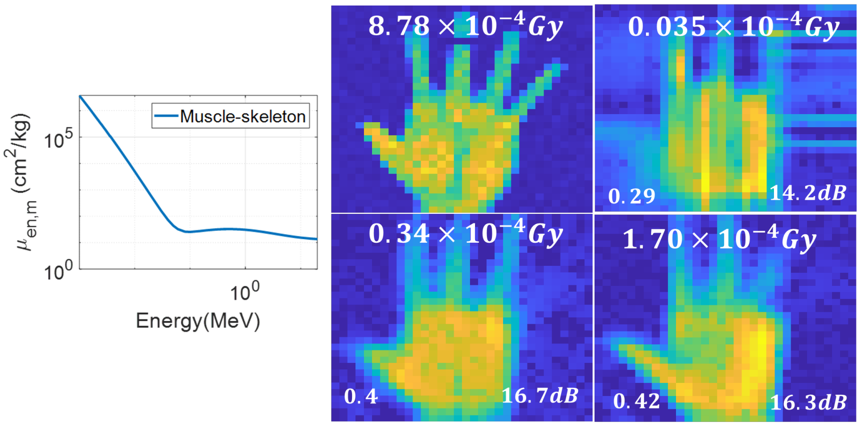

Figure 4.10: Left: Mass absorption coefficient $\mu_{en,m}$ in cm^2/kg as a function of incoming photons energies in MeV. Right: Ground-truth and reconstructed scenes for different levels of absorbed dose (top of each figure). The compression ratio remains fixed at 12.5%, and the dose is controlled with the transmittance per captured snapshot. The SSIM and PSNR for each scenario can be seen in the bottom left and right respectively. The BM3D is used in the denoising stage.

gun, the material is assumed to be *red brass*, a common alloy implemented in firearm manufacturing ($\rho = 8.44 g/cm^3$, 88% Cu, 10% Sn, 2% Zn). Figure 4.11-top shows the SSIM and the PSNR curves, for different code transmittances per shot, 0.1%, 1%, and 10%, as a function of the compression ratio. Notice that, even for a 0.1% of code transmittance, the algorithm shows good performance when the compression ratio is sufficient enough. Figure 4.11-bottom shows the SSIM and PSNR for 3 different compression ratios, 3.125%, 6.25% and 12.5%, as a function of the code transmittance per shot. It is clearly seen that, after a certain point, increasing the transmittance does not necessarily mean a better reconstruction in terms of SSIM and PSNR.

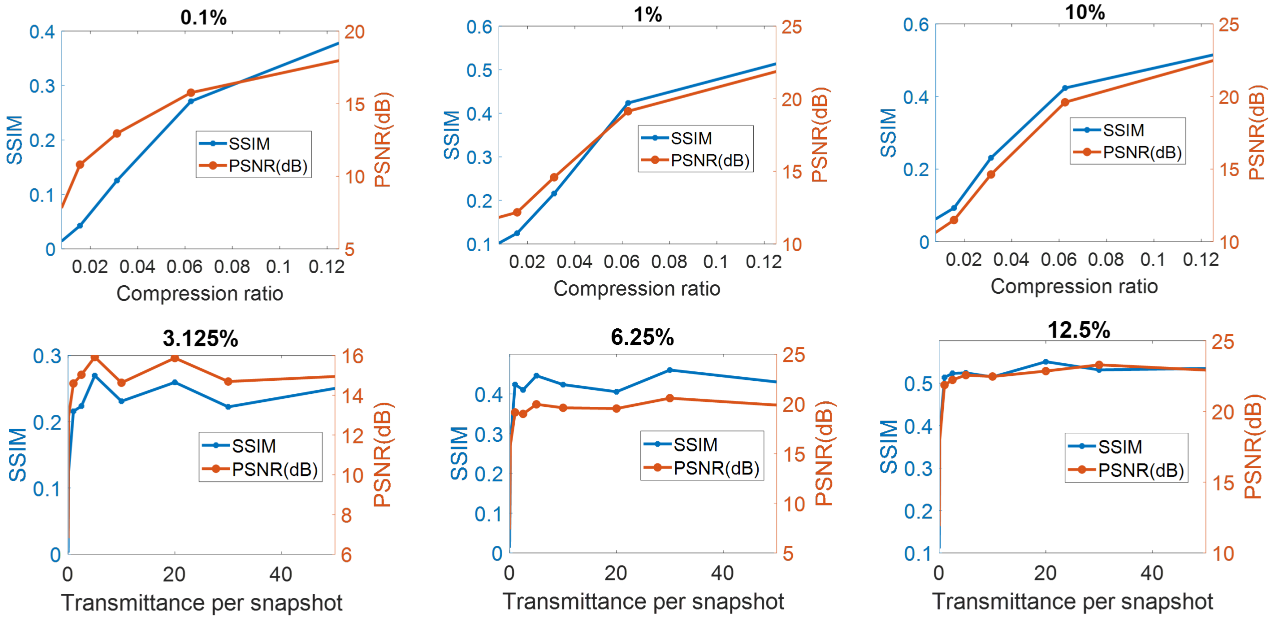

Figure 4.11: Top: SSIM and PSNR for different code transmittances per shot as a function of the compression ratio. From left to right: 0.1%, 1%, and 10% transmittance. Bottom: SSIM and PSNR for different compression ratios as a function of the code transmittance. From left to right: 3.125%, 6.25%, and 12.5% compression ratio. The BM3D was implemented as denoiser.

Results for 3 different coded transmittances, 1%, 2.5% and 5%, and a compression ratio of 12.5% can be seen in Fig. 4.12. As it can be seen, the firearm located in the body can be easily identified by a simple visual inspection; as stated in [66], in order to quantify the detectability of the suspected gun, a contrast estimation must be done; the contrast against the background is defined as follows:

$$C_{TR} = \left| \frac{N_{str} - N_{bkg}}{N_{bkg}} \right|, \tag{4.16}$$

where N_{str} and N_{bkg} are the mean values of the pixels on the target area and background, respectively. Fig. 4.12 shows a 22×22 pixel zoom over the image in the area

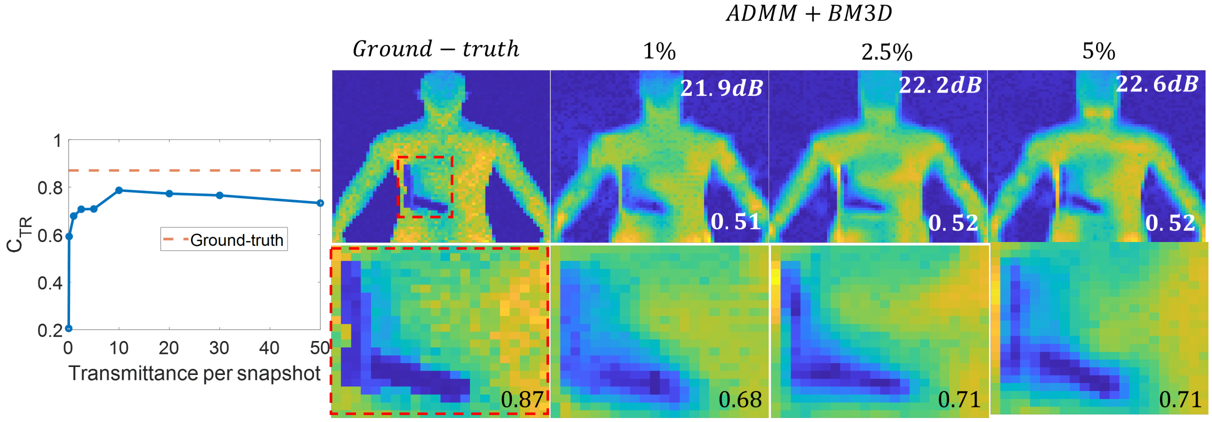

Figure 4.12: Left: Contrast against the background C_{TR} of the firearm as a function of the transmittance per snapshot for a fixed compression ratio of 12.5%. Top-right: Ground-truth and reconstructed scenes for three different transmittances per snapshot, with a fixed compression ratio of 12.5%. Bottom-right: Zooming portion of the fire-arm and its C_{TR} (bottom-right). The BM3D was implemented as denoiser.

where the firearm is placed, as well as the C_{TR} for each reconstruction (bottom-right). The background is assumed to be the human silhouette. The C_{TR} as a function of the code transmittance per snapshot can also be seen in Fig. 4.12. Notice how C_{TR} reaches a maximum when the transmittance is 10%.

4.9 Experimental demonstration

The CXBI was mounted in the Computational X-rays Imaging Laboratory at the University of Delaware, using off-the-shell hardware components. Figure 4.13 shows the test-bed implementation. A Thermo-fisher Micro CT X-ray source (PXS10) is powered at 130kV, with a current of $110\mu A$. Then, a lead collimator with a $1.6cm \times 1.6cm$ window is used in order to avoid unwanted scattering in the surroundings. The collimated radiation hits a tungsten-based mask, with pixel pitch size of $0.5mm \times 0.5mm$, and 50% transmittance, and then the coded radiation travels towards the water target. Finally, an X-ray detector (CareView 560-RF-DE) is used to capture the backscattered radiation. To assemble the sensing matrix $\mathbf{A}$, the target is first removed, and the detector is located in the target place. Then, as the coded aperture moves (using an automatized translational stage with speed of 0.8mm/s), each pattern that will be used in the reconstruction is registered.

After all the patterns are registered (384 in total), the target is placed in its original spot, and the detector moves backwards. Then, as the coded aperture moves, the backscattered radiation for each of its patterns is captured. After that, the target is removed, and the capturing process is repeated. This last step removes unwanted radiation in the measurements (see section 4.5.6)).

To assemble the sensing matrix, the coded aperture patterns need to be binarized. A grouping of 7×7 pixels is done, such that a coded aperture pitch size occupies a region of 7×7 pixels at the detector when running the calibration process. After that, a mean value that represents a single 7×7 region is taken; then, if that value surpasses a threshold, the pixel is set to 1, and if not, to 0. The threshold is set to be the mean value of all the grouped 7×7 regions. Figure 4.14 shows the original and binarized coded aperture for a given snapshot.

On the other side, each of the elements of the measurements vector $\mathbf{y}$ is defined by adding up all the pixel values of the detector measurements; this is because the CareView 560-RF-DE is a 2-dimensional detector; however, CXBI is done through a single pixel scintillator plate. Posteriorly, the measurements with no target are subtracted from the measurements with target, as it was done in the simulations. Then, the reconstruction algorithm is run. In order to avoid $\mathbf{y}$ having large values (the quantization bits of the sensor are 16), this vector is divided by 1×10^3 (see appendix D for more details). Results for the water target can be seen in Fig. 4.15-right; a projection of the target when the detector is located in the front part, is also shown as reference.

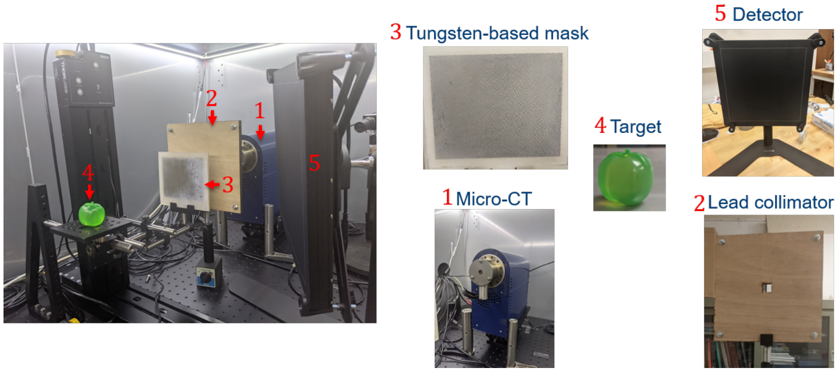

Figure 4.13: CXBI test-bed implementation. Five different hardware components can be observed. 1: Micro CT X-rays source. 2- Lead Collimator with a $1.6cm \times 1.6cm$ window. 3- Tungsten-based coded aperture with pitch size of $0.5mm \times 0.5mm$. 4- Water cube target. 5- Dual energy detector.

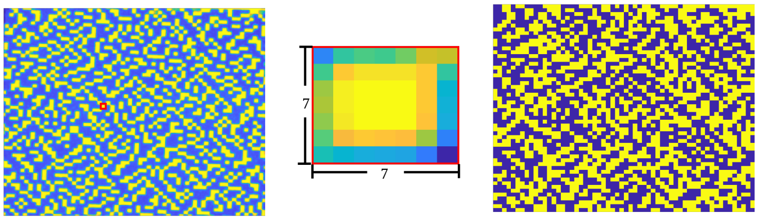

Figure 4.14: Left: Original captured pattern for a given snapshot. Middle: grouping of 7×7 pixels in the detector. Right: Binarized coded aperture used to assemble the sensing matrix.

4.10 Conclusions and ideas to further explore

This paper proposes the Compressive Compton X-rays Backscattering Imager (CXBI) as an alternative to the state-of-the-art scanning techniques. A discrete measurements model in accordance with a relativistic physics framework was proposed; this model assumes that the captured photons coming from the Compton effect are singly

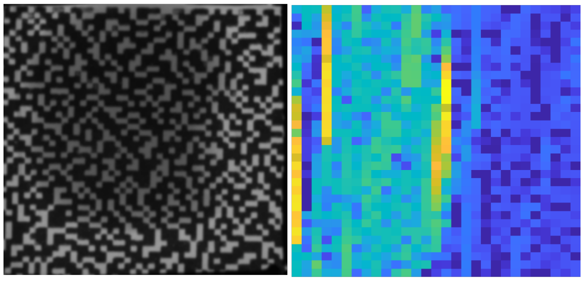

Figure 4.15: Left: forward projection of the target; the deep-gray area represents the target. Right: reconstructed scene of the water target using the CXBI experimental test-bed implementation.

scattered. A forward measurements model, based on the single pixel imaging principle, was developed; this model considers the movement of the coded aperture patterns, or equivalently the movement of the body while the coded aperture remains fixed. A functional CXBI experiment in GATE was programmed and tested for different conditions and bodies (see Fig. 4.5); a method to create the ground-truth images was also proposed and implemented. The viability of CXBI for human screening and the expected dose per pixel was analyzed for several scenarios, obtaining a good contrast estimation of the suspected body against the human silhouette. Finally, preliminary experimental results using off-the-shell hardware components are shown. Future work should focus on improving the experiments and reducing the needed dose to get accurate results.

Given the novelty of this work, a conference paper in *SPIE-Defense and Commercial Sensing* will be published in April 2022, and a journal paper in *Optics express* was submitted for publication and it is now under revision. All the data and software for this chapter are available in [96].

As future work, we propose to explore the CXBI with a two-dimensional detector. Given that the scintillator plates do not offer spatial information of the scene, this is a suboptimal approach that can be improved as depicted in Fig. 4.16. Also, we propose to extend the CXBI to estimate the depth; given that the cross-sectional area involved in the scattering process depends on the relative position of the beam with respect to the detector, this can be used as a parameter to estimate the depth from which singly scattered photons come, as depicted in Fig. 4.17 (see Dinca et al. [68] and Asad et al. [97]).

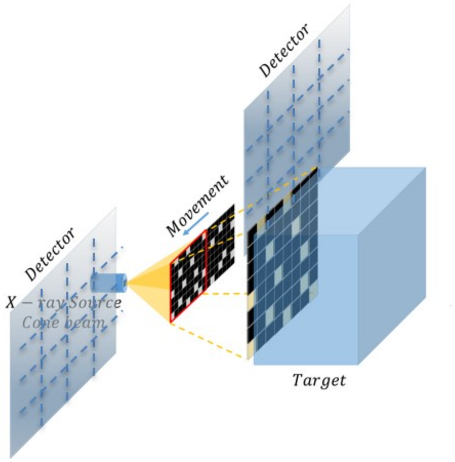

Figure 4.16: Proposed CXBI implementation using two-dimensional detectors.

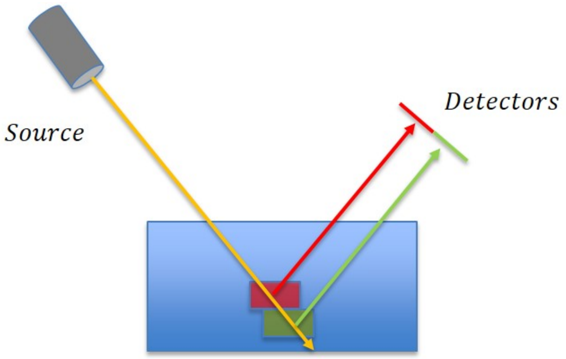

Figure 4.17: Use of Compton Backscattered photons to determine the depth (already studied in the literature). We propose to extend CXBI for depth estimation based on this principle.

BIBLIOGRAPHY

[1] RH Dicke. Scatter-hole cameras for x-rays and gamma rays. *The astrophysical journal*, 153:L101, 1968.

[2] Stephen R Gottesman and E Edward Fenimore. New family of binary arrays for coded aperture imaging. *Applied optics*, 28(20):4344–4352, 1989.

[3] Richard G Baraniuk. Compressive sensing [lecture notes]. *IEEE signal processing magazine*, 24(4):118–121, 2007.

[4] Guolan Lu and Baowei Fei. Medical hyperspectral imaging: a review. *Journal of biomedical optics*, 19(1):010901, 2014.

[5] Dimitris G Manolakis, Ronald B Lockwood, and Thomas W Cooley. *Hyperspectral imaging remote sensing: physics, sensors, and algorithms.* Cambridge University Press, 2016.

[6] Haida Liang. Advances in multispectral and hyperspectral imaging for archaeology and art conservation. *Applied Physics A*, 106(2):309–323, 2012.

[7] Yuwei Liu, Hongbin Pu, and Da-Wen Sun. Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications. *Trends in food science & technology*, 69:25–35, 2017.

[8] David J Brady. *Optical imaging and spectroscopy.* John Wiley & Sons, 2009.

[9] Ashwin Wagadarikar, Renu John, Rebecca Willett, and David Brady. Single disperser design for coded aperture snapshot spectral imaging. *Applied optics*, 47(10):B44–B51, 2008.

[10] Laura Galvis-Carreno. *Coded aperture design in compressive spectral imaging.* University of Delaware, 2018.

[11] Yuehao Wu, Iftekhar O Mirza, Gonzalo R Arce, and Dennis W Prather. Development of a digital-micromirror-device-based multishot snapshot spectral imaging system. *Optics letters*, 36(14):2692–2694, 2011.

[12] Henry Arguello and Gonzalo R Arce. Colored coded aperture design by concentration of measure in compressive spectral imaging. *IEEE Transactions on Image Processing*, 23(4):1896–1908, 2014.

[13] Henry Arguello, Hoover Rueda, Yuehao Wu, Dennis W Prather, and Gonzalo R Arce. Higher-order computational model for coded aperture spectral imaging. *Applied optics*, 52(10):D12–D21, 2013.

[14] Laura Galvis, Henry Arguello, Daniel Lau, and Gonzalo R Arce. Side information in coded aperture compressive spectral imaging. In *Emerging Digital Micromirror Device Based Systems and Applications IX*, volume 10117, page 101170H. International Society for Optics and Photonics, 2017.

[15] Miguel Marquez, Hoover Rueda-Chacon, and Henry Arguello. Compressive spectral imaging via virtual side information. *IEEE Transactions on Computational Imaging*, 7:114–123, 2021.

[16] Claudia V Correa, Henry Arguello, and Gonzalo R Arce. Snapshot colored compressive spectral imager. *JOSA A*, 32(10):1754–1763, 2015.

[17] Xing Lin, Yebin Liu, Jiamin Wu, and Qionghai Dai. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Transactions on Graphics (TOG)*, 33(6):1–11, 2014.

[18] Michael E Gehm, Renu John, David J Brady, Rebecca M Willett, and Timothy J Schulz. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics express*, 15(21):14013–14027, 2007.

[19] Alex Chalmers. Applications of backscatter x-ray imaging sensors for homeland defense. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Defense and Law Enforcement II*, volume 5071, pages 388–396. International Society for Optics and Photonics, 2003.

[20] MM Hosamani and NM Badiger. Determination of effective atomic number of composite materials using backscattered gamma photons–a novel method. *Chemical Physics Letters*, 695:94–98, 2018.

[21] Alex Chalmers. Rapid inspection of cargos at portals using drive-through transmission and backscatter x-ray imaging. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense III*, volume 5403, pages 644–648. International Society for Optics and Photonics, 2004.

[22] Daniel Shedlock. *X-ray backscatter imaging for radiography by selective detection and snapshot: Evolution, development, and optimization.* 2007.

[23] Jie Xu, Xin Wang, Qi Zhan, Shengling Huang, Yifan Chen, and Baozhong Mu. A novel lobster-eye imaging system based on schmidt-type objective for x-ray-backscattering inspection. *Review of Scientific Instruments*, 87(7):073103, 2016.

[24] Anthony A Faust, Richard E Rothschild, Philippe Leblanc, and John Elton McFee. Development of a coded aperture x-ray backscatter imager for explosive device detection. *IEEE Transactions on nuclear science*, 56(1):299–307, 2009.

[25] Takayoshi Shimura, Takuji Hosoi, and Heiji Watanabe. Backscattering x-ray imaging using fresnel zone aperture. *Applied Physics Express*, 14(7):072002, 2021.

[26] David J Brenner, Richard Doll, Dudley T Goodhead, Eric J Hall, Charles E Land, John B Little, Jay H Lubin, Dale L Preston, R Julian Preston, Jerome S Puskin, et al. Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. *Proceedings of the National Academy of Sciences*, 100(24):13761–13766, 2003.

[27] David J Brenner. Are x-ray backscatter scanners safe for airport passenger screening? for most individuals, probably yes, but a billion scans per year raises long-term public health concerns. *Radiology*, 259(1):6–10, 2011.

[28] Xun Cao, Tao Yue, Xing Lin, Stephen Lin, Xin Yuan, Qionghai Dai, Lawrence Carin, and David J Brady. Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world. *IEEE Signal Processing Magazine*, 33(5):95–108, 2016.

[29] Ankit Mohan, Ramesh Raskar, and Jack Tumblin. Agile spectrum imaging: Programmable wavelength modulation for cameras and projectors. In *Computer Graphics Forum*, volume 27, pages 709–717. Wiley Online Library, 2008.

[30] Henry Arguello, Hoover F Rueda, and Gonzalo R Arce. Spatial super-resolution in code aperture spectral imaging. In *Compressive Sensing*, volume 8365, page 83650A. International Society for Optics and Photonics, 2012.

[31] Hoover F Rueda, Henry Arguello, and Gonzalo R Arce. On super-resolved coded aperture spectral imaging. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX*, volume 8743, page 87431B. International Society for Optics and Photonics, 2013.

[32] Gonzalo R Arce, David J Brady, Lawrence Carin, Henry Arguello, and David S Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, 2013.

[33] Mário AT Figueiredo, Robert D Nowak, and Stephen J Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of selected topics in signal processing*, 1(4):586–597, 2007.

[34] Marco F Duarte and Richard G Baraniuk. Kronecker compressive sensing. *IEEE Transactions on Image Processing*, 21(2):494–504, 2011.

[35] Edgar Salazar, Alejandro Parada, and Gonzalo R Arce. Spatial super-resolution reconstruction via sscsi compressive spectral imagers. In *Computational Optical Sensing and Imaging*, pages CTu5D–5. Optical Society of America, 2018.

[36] Edgar Salazar, Alejandro Parada, and Gonzalo R Arce. Spectral zooming in sscsi compressive spectral imagers. In *Computational Optical Sensing and Imaging*, pages CTu5D–7. Optical Society of America, 2018.

[37] Edgar Salazar, Alejandro Parada-Mayorga, and Gonzalo R Arce. Spectral zooming and resolution limits of spatial spectral compressive spectral imagers. *IEEE Transactions on Computational Imaging*, 5(2):165–179, 2019.

[38] SSCSI-Sensing Model Documentation. `https://github.com/Edgar-Noita/SSCSI-Sensing-Matrix-Code`.

[39] Xin Yuan, David J Brady, and Aggelos K Katsaggelos. Snapshot compressive imaging: Theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 38(2):65–88, 2021.

[40] Emmanuel J Candes. The restricted isometry property and its implications for compressed sensing. *Comptes rendus mathematique*, 346(9-10):589–592, 2008.

[41] Ashwin Ashok Wagadarikar. *Compressive spectral and coherence imaging*. Duke University, 2010.

[42] Claudia V Correa, Henry Arguello, and Gonzalo R Arce. Spatiotemporal blue noise coded aperture design for multi-shot compressive spectral imaging. *JOSA A*, 33(12):2312–2322, 2016.

[43] Henry Arguello and Gonzalo R Arce. Rank minimization code aperture design for spectrally selective compressive imaging. *IEEE transactions on image processing*, 22(3):941–954, 2012.

[44] Hoover Rueda, Henry Arguello, and Gonzalo R Arce. High-dimensional optimization of color coded apertures for compressive spectral cameras. In *2017 25th European Signal Processing Conference (EUSIPCO)*, pages 463–467. IEEE, 2017.

[45] Kareth M León-López, Laura V Galvis Carreno, and Henry Arguello Fuentes. Temporal colored coded aperture design in compressive spectral video sensing. *IEEE Transactions on Image Processing*, 28(1):253–264, 2018.

[46] Lizhi Wang, Tao Zhang, Ying Fu, and Hua Huang. Hyperreconnet: Joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging. *IEEE Transactions on Image Processing*, 28(5):2257–2270, 2018.

[47] Jorge Bacca, Laura Galvis, and Henry Arguello. Coupled deep learning coded aperture design for compressive image classification. *Optics express*, 28(6):8528–8540, 2020.

[48] Andreas M Tillmann and Marc E Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Transactions on Information Theory*, 60(2):1248–1259, 2013.

[49] Wen Hou and Cishen Zhang. Analysis of compressed sensing based ct reconstruction with low radiation. In *2014 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pages 291–296. IEEE, 2014.

[50] Angela P Cuadros and Gonzalo R Arce. Coded aperture optimization in compressive x-ray tomography: a gradient descent approach. *Optics express*, 25(20):23833–23849, 2017.

[51] Alejandro Parada-Mayorga and Gonzalo R Arce. Colored coded aperture design in compressive spectral imaging via minimum coherence. *IEEE Transactions on Computational Imaging*, 3(2):202–216, 2017.

[52] Hoover Rueda-Chacon, Juan F Florez-Ospina, Daniel L Lau, and Gonzalo R Arce. Snapshot compressive tof+ spectral imaging via optimized color-coded apertures. *IEEE transactions on pattern analysis and machine intelligence*, 42(10):2346–2360, 2019.

[53] Claudia V Correa, Henry Arguello, and Gonzalo R Arce. Testbed implementation of a compressive spectral imaging system with spatio temporal blue noise coded apertures. In *Computational Optical Sensing and Imaging*, pages CW5D–8. Optical Society of America, 2016.

[54] Nelson Diaz, Carlos Hinojosa, and Henry Arguello. Adaptive grayscale compressive spectral imaging using optimal blue noise coding patterns. *Optics & Laser Technology*, 117:147–157, 2019.

[55] Hao Zhang, Xu Ma, Daniel L Lau, Jianchen Zhu, and Gonzalo R Arce. Compressive spectral imaging based on hexagonal blue noise coded apertures. *IEEE Transactions on Computational Imaging*, 6:749–763, 2020.

[56] Robert A Ulichney. Dithering with blue noise. *Proceedings of the IEEE*, 76(1):56–79, 1988.

[57] Daniel L Lau, Gonzalo R Arce, and Neal C Gallagher. Green-noise digital halftoning. *Proceedings of the IEEE*, 86(12):2424–2444, 1998.

[58] Daniel L Lau, Gonzalo R Arce, and Neal C Gallagher. Digital halftoning by means of green-noise masks. *JOSA A*, 16(7):1575–1586, 1999.

[59] Daniel L Lau, Gonzalo R Arce, and Neal C Gallagher. Digital color halftoning with generalized error diffusion and multichannel green-noise masks. *IEEE Transactions on Image Processing*, 9(5):923–935, 2000.

[60] Daniel L Lau and Gonzalo R Arce. *Modern digital halftoning*, volume 1. CRC Press, 2018.

[61] CAVE multispectral database (accessed: 2019-02-09). `http://www.cs.columbia.edu/CAVE/databases/multispectral`.

[62] Edgar Salazar and Gonzalo R Arce. Optimal coding patterns in spatial spectral compressive spectral imagers. In *Propagation Through and Characterization of Atmospheric and Oceanic Phenomena*, pages JW4B–3. Optical Society of America, 2019.

[63] Edgar Salazar and Gonzalo R Arce. Coded aperture optimization in spatial spectral compressive spectral imagers. *IEEE Transactions on Computational Imaging*, 6:764–777, 2020.

[64] Pan Wang, Jie Li, Chun Qi, Lin Wang, and Jieru Chen. Input aperture restriction of the spatial spectral compressive spectral imager and a comprehensive solution for it. *Optics Express*, 29(12):17875–17889, 2021.

[65] SSCSI-Coded Aperture Optimization Documentation. `https://github.com/Edgar-Noita/SSCSI-Coded-aperture-optimization`.

[66] Zongjian Cao. Optimization for the tradeoff of detection efficiency and absorbed dose in x-ray backscatter imaging. *Journal of Transportation Security*, 6(1):59–76, 2013.

[67] Baolu Yang, Xin Wang, Haiping Shen, Jie Xu, Kai Xiong, and Baozhong Mu. Design of x-ray backscatter imaging system for vehicle detection. In *Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies IV*, volume 11542, page 115420L. International Society for Optics and Photonics, 2020.

[68] Dan-Cristian Dinca, Jeffrey R Schubert, and J Callerame. X-ray backscatter imaging. In *Optics and Photonics in Global Homeland Security IV*, volume 6945, page 694516. International Society for Optics and Photonics, 2008.

[69] Shengling Huang, Xin Wang, Yifan Chen, Jie Xu, and Baozhong Mu. Simulation on x-rays backscatter imaging based on monte carlo methods for security inspection. In *Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies II*, volume 10802, page 1080203. International Society for Optics and Photonics, 2018.

[70] Shengling Huang, Xin Wang, Yifan Chen, Jie Xu, Tian Tang, and Baozhong Mu. Modeling and quantitative analysis of x-ray transmission and backscatter imaging aimed at security inspection. *Optics express*, 27(2):337–349, 2019.

[71] Arthur H Compton. A quantum theory of the scattering of x-rays by light elements. *Physical review*, 21(5):483, 1923.

[72] John Howard Hubbell, Wm J Veigele, EA Briggs, RT Brown, DT Cromer, and d RJ Howerton. Atomic form factors, incoherent scattering functions, and photon scattering cross sections. *Journal of physical and chemical reference data*, 4(3):471–538, 1975.

[73] Stephen Boyd, Neal Parikh, and Eric Chu. *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.

[74] Siming Zheng, Yang Liu, Ziyi Meng, Mu Qiao, Zhishen Tong, Xiaoyu Yang, Shensheng Han, and Xin Yuan. Deep plug-and-play priors for spectral snapshot compressive imaging. *Photonics Research*, 9(2):B18–B29, 2021.

[75] Vishal Monga, Yuelong Li, and Yonina C Eldar. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*, 38(2):18–44, 2021.

[76] Catherine F Higham, Roderick Murray-Smith, Miles J Padgett, and Matthew P Edgar. Deep learning for real-time single-pixel video. *Scientific reports*, 8(1):1–9, 2018.

[77] Ikuo Hoshi, Tomoyoshi Shimobaba, Takashi Kakue, and Tomoyoshi Ito. Single-pixel imaging using a recurrent neural network combined with convolutional layers. *Optics Express*, 28(23):34069–34078, 2020.

[78] Nazmul Karim and Nazanin Rahnavard. Spi-gan: Towards single-pixel imaging through generative adversarial network. *arXiv preprint arXiv:2107.01330*, 2021.

[79] G Rigaud and BN Hahn. Reconstruction algorithm for 3d compton scattering imaging with incomplete data. *Inverse Problems in Science and Engineering*, 29(7):967–989, 2021.

[80] James W Webber and William RB Lionheart. Three dimensional compton scattering tomography. *Inverse Problems*, 34(8):084001, 2018.

[81] Joel A Greenberg, Mehadi Hassan, Kalyani Krishnamurthy, and David Brady. Structured illumination for tomographic x-ray diffraction imaging. *Analyst*, 139(4):709–713, 2014.

[82] Zheyuan Zhu, Ryan A Ellis, and Shuo Pang. Coded cone-beam x-ray diffraction tomography with a low-brilliance tabletop source. *Optica*, 5(6):733–738, 2018.

[83] Audrey Kueh, Jason M Warnett, Gregory J Gibbons, Julia Brettschneider, Thomas E Nichols, Mark A Williams, and Wilfrid S Kendall. Modelling the penumbra in computed tomography 1. *Journal of X-ray Science and Technology*, 24(4):583–597, 2016.

[84] Marco F Duarte, Mark A Davenport, Dharmpal Takhar, Jason N Laska, Ting Sun, Kevin F Kelly, and Richard G Baraniuk. Single-pixel imaging via compressive sampling. *IEEE signal processing magazine*, 25(2):83–91, 2008.

[85] Stephen J Norton. Compton scattering tomography. *Journal of applied physics*, 76(4):2007–2015, 1994.

[86] Otha W Linton and Fred A Mettler Jr. National council on radiation protection and measurements. In *National conference on dose reduction in CT, with an emphasis on pediatric patients. AJR Am J Roentgenol*, volume 181, pages 321–329, 2003.

[87] Jacob Punnoose, J Xu, A Sisniega, W Zbijewski, and JH Siewerdsen. spektr 3.0, a computational tool for x-ray spectrum modeling and analysis. *Medical physics*, 43(8Part1):4711–4717, 2016.

[88] Peter Rez, Robert L Metzger, and Kenneth L Mossman. The dose from compton backscatter screening. *Radiation protection dosimetry*, 145(1):75–81, 2011.

[89] Jessica C Hsu, Lenitza M Nieves, Oshra Betzer, Tamar Sadan, Peter B Noël, Rachela Popovtzer, and David P Cormode. Nanoparticle contrast agents for x-ray imaging applications. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 12(6):e1642, 2020.

[90] Rolf Behling. *Modern diagnostic x-ray sources: technology, manufacturing, reliability*. CRC Press, 2021.

[91] Gate Documentation. `https://opengate.readthedocs.io`. Accessed: 2021-07-16.

[92] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.

[93] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Transactions on Image Processing*, 27(9):4608–4622, 2018.

[94] Tianyi Mao, Angela P Cuadros, Xu Ma, Weiji He, Qian Chen, and Gonzalo R Arce. Coded aperture optimization in x-ray tomography via sparse principal component analysis. *IEEE Transactions on Computational Imaging*, 6:73–86, 2019.

[95] JH Hubbell and SM Seltzer. Nist standard reference database 126. *Gaithersburg, MD: National Institute of Standards and Technology*, 1996.

[96] CXBI-Documentation. `https://github.com/Edgar-Noita/CXBI`.

[97] Z Asa'd, M Asghar, and DC Imrie. The measurement of the wall thickness of steel sections using compton backscattering. *Measurement Science and Technology*, 8(4):377, 1997.

Appendix A

CALCULATION OF THE W PARAMETERS IMPLEMENTED IN THE SSCSI DISCRETE MEASUREMENTS MODEL

As mentioned in chapter 2 , $W_{m,m'}$ indicates the fraction of the coded aperture $\mathbf{t}_{m',n'}$ impinging on the sensor element (m,n). A graphical description of $W_{m,m'}$ can be seen in Fig. A.1. As depicted, three cases can be distinguished:

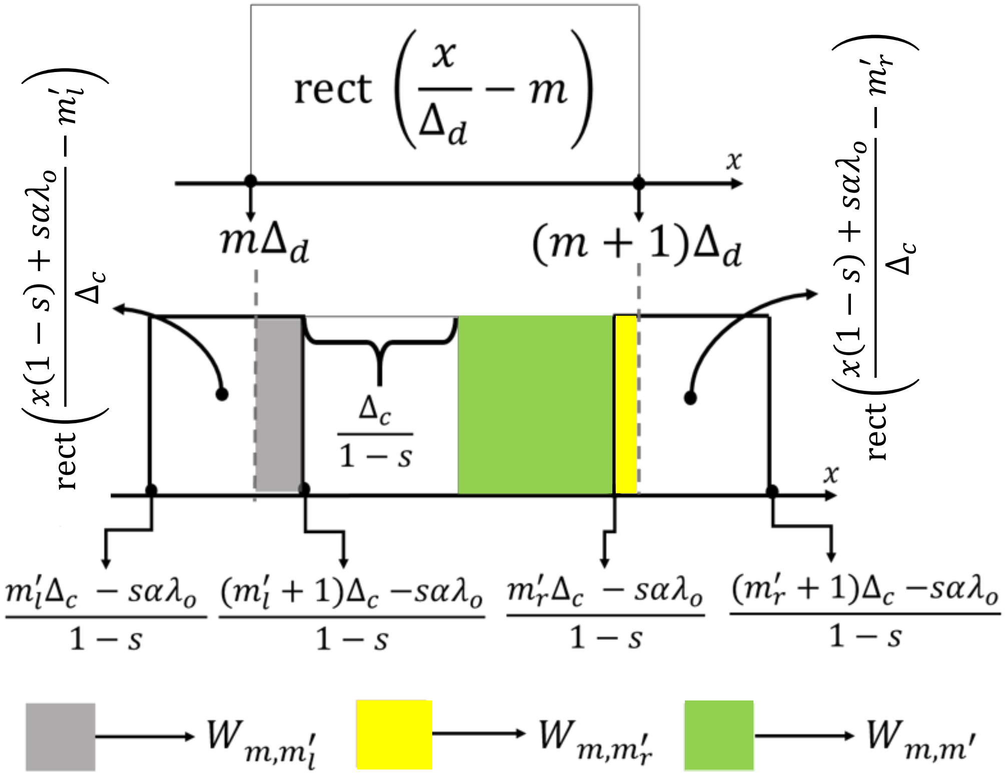

Figure A.1: Graphical explanation to calculate W parameters implemented in Eq. 2.11.

Gray Area

When located at the gray area, $m' = m'_l$, and W_{m,m'_l} can be calculated as

$$W_{m,m'_l} = \frac{\frac{(m'_l+1)\Delta_c - s\alpha\lambda_o}{1-s} - (m)\Delta_d}{\Delta_c/(1-s)}. \tag{A.1}$$

The term $\Delta_c/(1-s)$ is used as a normalization parameter that allows one to obtain the percentage of $\mathbf{t}_{m'_l,n'}$ measured by the sensor element (m, n).

Yellow Area

When located at the yellow area, $m' = m'_r$, and W_{m,m'_r} can be calculated as

$$W_{m,m'_r} = \frac{(m+1)\Delta_d - \frac{m'_r\Delta_c - s\alpha\lambda_o}{1-s}}{\Delta_c/(1-s)}. \tag{A.2}$$

Green Area

When located at the green area, $m'_l < m' < m'_r$, and $W_{m,m'}$ can be calculated as

$$W_{m,m'} = \frac{\Delta_c/(1-s)}{\Delta_c/(1-s)} = 1 \tag{A.3}$$

The fact $W_{m,m'} = 1$ indicates that $\mathbf{t}_{m',n'}$ is fully impinging on sensor element (m, n). Equation (2.12) contains a general expression for $W_{m,m'}$ considering all the mentioned cases. An example of how to calculate $W_{m,m'}$ can be done based on Fig. 2.6 right. In that figure, $W_{0,2} = \frac{\Delta_d - \frac{2\Delta_c}{(1-s)}}{\Delta_c/(1-s)}$, $W_{1,2} = \frac{\frac{3\Delta_c}{(1-s)} - \Delta_d}{\Delta_c/(1-s)}$ and $W_{0,1} = 1$.

Appendix B

CALCULATION OF THE SYNTHETIC CODED APERTURES IMPLEMENTED IN THE SSCSI DISCRETE MEASUREMENTS MODEL

The effective coded aperture $\tilde{\mathbf{t}}_{m,n'}$ in Eq. (2.13) is given by the expression $\tilde{\mathbf{t}}_{m,n'} = \mathbf{t}_{m'-1,n'} \times p_m + \mathbf{t}_{m',n'} \times (1 - p_m)$, where n' is given by Eq. (2.9), $m' = \left\lfloor \frac{(m)\Delta_d(1-s)+s\alpha\lambda_o}{\Delta_c} \right\rfloor + 1$ and the parameter p_m indicates the percentage of the $(m,n)^{th}$ sensor element occupied by the $(m'-1,n')^{th}$ coded aperture element. Figure B.1 contains a graphical explanation to calculate p_m. Notice that, according to this figure, $m\Delta_d \leq \frac{m'\Delta_c - s\alpha\lambda_o}{1-s} \leq (m+1)\Delta_d$, or equivalently $m' = \left\lfloor \frac{(m)\Delta_d(1-s)+s\alpha\lambda_o}{\Delta_c} \right\rfloor + 1$ (considering that m' is an integer index).

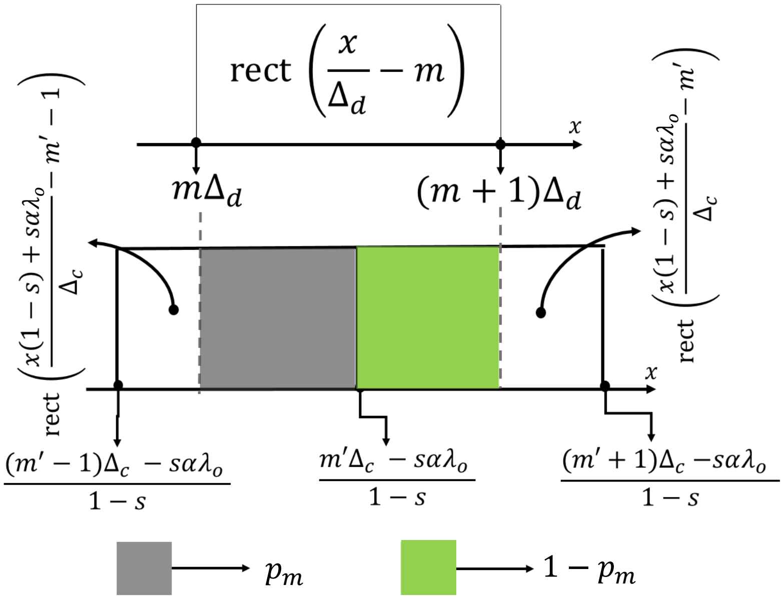

Figure B.1: Graphical explanation to calculate the variable p_m in Eq. 2.13.

The portion of the $(m, n)^{th}$ sensor element occupied by $\mathbf{t}_{m'-1,n'}$ (gray area in Fig. B.1) can be calculated as

$$p_m = \frac{\frac{m'\Delta_c - s\alpha\lambda_o}{(1-s)} - (m)\Delta_d}{\Delta_d}, \tag{B.1}$$

where the sensor pitch size Δ_d appears in the denominator for normalization purposes. If $\frac{m'\Delta_c - s\alpha\lambda_o}{1-s} \geq (m+1)\Delta_d$, $\mathbf{t}_{m'-1,n'}$ fully occupies the $(m, n)^{th}$ sensor element and therefore $p_m = 1$. A general expression to calculate p_m, including the two cases mentioned can be seen in Eq. (2.14).

Appendix C

APPENDIX A: CALCULATION OF SUB-INDEXES IMPLEMENTED IN EQ. (3.13)

Let $j = \left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i$, $m_i = \left\lfloor \frac{i-1}{N} \right\rfloor - \left\lfloor \frac{i-1}{N^2} \right\rfloor N$, $n_i = \mathrm{mod}(i-1, N$ and $\ell_i = \left\lfloor \frac{i-1}{N^2} \right\rfloor + 1$. In order to calculate an expression for m_j, n_j and ℓ_j, the expression for j must be directly replaced into the expressions for m_i, n_i and k_i, as follows:

$$\begin{aligned} m_j &= \left\lfloor \frac{\left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i - 1}{N} \right\rfloor - \left\lfloor \frac{\left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i - 1}{N^2} \right\rfloor N \\ n_j &= \mathrm{mod}\left(\left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i - 1, N\right) \\ \ell_j &= \left\lfloor \frac{\left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N^2 + i - 1}{N^2} \right\rfloor . \end{aligned} \tag{C.1}$$

The last expression can be rewritten as follows

$$\begin{aligned} m_j &= \left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N + \left\lfloor \frac{i-1}{N} \right\rfloor - \left(k - \left\lfloor \frac{i-1}{N^2} \right\rfloor\right) N - \left\lfloor \frac{i-1}{N^2} \right\rfloor N \\ n_j &= \mathrm{mod}\left(i-1, N\right) \\ \ell_j &= k - \left\lfloor \frac{i-1}{N^2} \right\rfloor + \left\lfloor \frac{i-1}{N^2} \right\rfloor . \end{aligned} \tag{C.2}$$

Therefore, $m_j = m_i$, $n_j = n_i$ and $\ell_j = k$. Notice that, if one needs to evaluate the inner product of $\langle \mathbf{h}_i, \mathbf{h}_j \rangle$, one needs to take into account the coded aperture elements over a given row of the pattern n_i, on a neighborhood of length L.

Appendix D

EXPERIMENTAL DATA FOR CXBI

This appendix shows the experimental data captured in the laboratory to prove the functionality of the CXBI. All the available data can be found in [96]. Figure D.1-left shows the first four captured patterns using the two-dimensional detector in a forward position, while the binarized patterns after the thresholding can be seen in Fig. D.1-right.

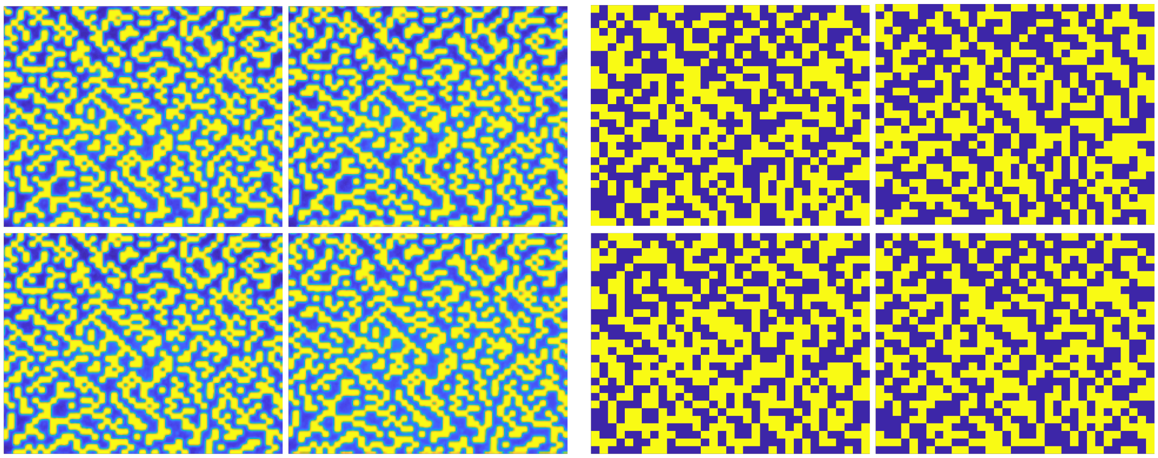

Figure D.1: Left: First 4 patterns captured in the CXBI calibration process. Right: Binary version of the patterns.

The measurements vector $\mathbf{y}$ with target, no target, and the resultant vector after the subtraction can be seen in Fig. D.2.

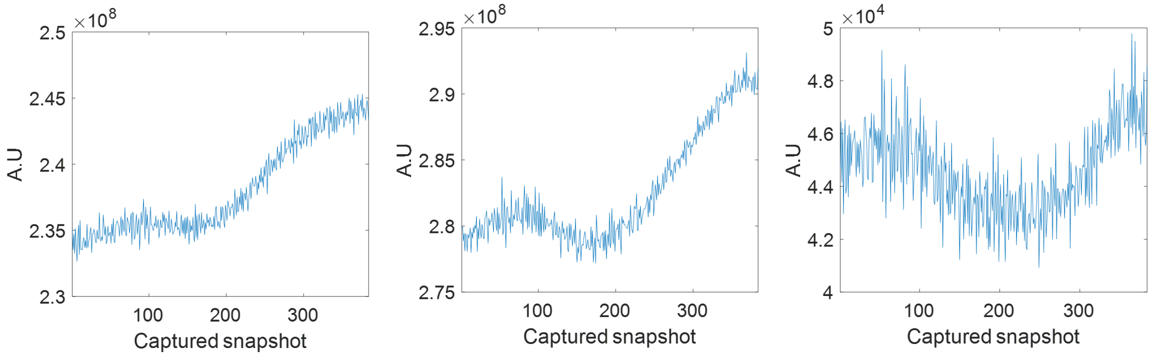

Figure D.2: Left: Measurements vector $\mathbf{y}$ with no target. Middle: Measurements vector $\mathbf{y}$ with target. Right: Subtraction of vectors to reduce the air and surroundings scattering. This vector is divided by 1×10^3 for convergence purposes of the reconstruction algorithm.

Appendix E

PERMISSIONS

This appendix contains the copyright notice for the material used in this thesis:

- Chapter 2 was published in IEEE transactions on Computational Imaging © 2019 IEEE. Reprinted with permission of authors, Edgar Salazar, Alejandro Parada-Mayorga, and Gonzalo Arce. Some of the results were also published in Computational, Optics, Sensing and Imaging (COSI), © Optica 2018.
- Chapter 3 was published in IEEE transactions on Computational Imaging © 2020 IEEE. Reprinted with permission of authors, Edgar Salazar, and Gonzalo Arce. Some of the results were also published in Computational, Optics, Sensing and Imaging (COSI), © Optica 2019.
- Part of chapter 4 was published in SPIE-Defense and Commercial Sensing April-2022 © SPIE 2022. Reprinted with permission of authors, Edgar Salazar, and Gonzalo Arce. Part of Chapter 4 was accepted for publication in Optics-express © Optica 2022. Reprinted with permission of authors, Edgar Salazar, Xiaokang Liu, and Gonzalo Arce.

In reference to IEEE copyrighted material which is used with permission in this thesis: the IEEE does not endorse any of University of Delaware products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to `https://www.ieee.org/publications/rights/rights-link.html` to learn how to obtain a License from RightsLink.

In reference to Optica copyrighted material which is used with permission in this thesis: One print or electronic copy of these papers may be made for personal use only. Systematic reproduction and distribution, duplication of any material in these papers for a fee or for commercial purposes, or modifications of the content of these papers are prohibited. For further information regarding Optica copyright and licensing information please refer to `https://opg.optica.org/submit/review/copyright_permissions.cfm#posting`.

In reference to SPIE copyrighted material which is used with permission in this thesis: Copyright 2022, Society of PhotoOptical Instrumentation Engineers (SPIE). One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this publication for a fee or for commercial purposes, and modification of the contents of the publication are prohibited. For further information regarding SPIE copyright and licensing information please refer to `https://spie.org/publications/contact-spie-publications`.

Some of the objects used in chapter 2 and 3 were not endorsed by the trademark owners and they are used here to illustrate the quality of SSCSI reconstructions. LEGO is a trademark of the LEGO Group, which does not sponsor, authorize or endorse the images in this dissertation. The LEGO Group. All Rights Reserved `http://aboutus.lego.com/en-us/legal-notice/fair-play/`

Some of the STL files implemented in the GATE simulations were downloaded from `https://www.cgtrader.com/` and `https://cults3d.com`. Free access.